

Wed Mar 12 10:08:37 2003

us-10-046-433-40.rsp

Page 1

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2003, 23:10:30 ; Search time 23 Seconds
(without alignments)
1805.122 Million cell updates/sec

Title: US-10-046-433-40

Perfect score: 5506
Sequence: 1 MAEPGSHHLSARVGRTER.....LGRSNHLPRLGLMDLTQCR 1001

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	251	4.6	1609	1	LMG1_HUMAN
2	225	4.1	1607	1	LMG1_MOUSE
3	224.5	4.1	1639	1	LMG1_DROME
4	211.5	3.8	713	1	TSN4_GIALA
5	209.5	3.8	1581	1	LMG3_MOUSE
6	208.5	3.8	1877	1	PKC5_MOUSE
7	208.5	3.8	1696	1	PKC5_BRACL
8	193	3.5	687	1	VSAL1_GIALA
9	192	3.5	1557	1	LMG3_HUMAN
10	192	3.5	1587	1	LMG1_HUMAN
11	174	3.2	1786	1	LMG1_MOUSE
12	172.5	3.1	3106	1	TS11_GIALA
13	171.5	3.1	1786	1	LMG1_MOUSE
14	170	3.1	1680	1	FUR2_DROME
15	169	3.1	2768	1	THYG_RAT
16	168.5	3.1	2768	1	LMG2_HUMAN
17	167	3.0	610	1	LMG2_HUMAN
18	166.5	3.0	3110	1	LMG1_MOUSE
19	163	3.0	3084	1	LMG1_MOUSE
20	162	2.9	1246	1	YMW2_CAEEL
21	161.5	2.9	2524	1	NOTC_XENLA
22	160.5	2.9	1790	1	LMG1_DROME
23	160.5	2.9	3672	1	LMG2_CAEEL
24	160	2.9	993	1	EPB3_MOUSE
25	160	2.9	2470	1	NCG2_MOUSE
26	159	2.9	1169	1	YK82_YEAST
27	158	2.9	1218	1	JAG1_HUMAN
28	158	2.9	2871	1	FRN1_BOVIN
29	158	2.9	3712	1	LMG2_HUMAN
30	157.5	2.9	1193	1	LMG2_HUMAN
31	157.5	2.9	2907	1	EPB2_MOUSE
32	157.5	2.9	2911	1	FBN2_HUMAN
33	157.5	2.9	3075	1	LMG1_HUMAN

34	157	2.9	4393	1	PCBM_HUMAN
35	156	2.8	1219	1	JAG1_RAT
36	155.5	2.8	1242	1	JAG1_BRARE
37	155	2.8	998	1	EPB3_HUMAN
38	155	2.8	1218	1	EPB3_MOUSE
39	154.5	2.8	4544	1	JAG1_MOUSE
40	154	2.8	2471	1	LMG1_HUMAN
41	154	2.8	2766	1	THYG_MOUSE
42	154	2.8	2871	1	FRN1_MOUSE
43	152.5	2.8	974	1	EPB3_XENLA
44	151	2.7	3695	1	LMG5_HUMAN
45	151	2.7	5376	1	ZAN_MOUSE

ALIGNMENTS

RESULT 1	LMG1_HUMAN	STANDARD;	PRT;	1609 AA.
ID	LMG1_HUMAN			
AC	P11047;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-NOV-1991 (Rel. 20, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Laminin gamma-1 chain precursor (Laminin B2 chain).			
GN	LMG1 OR LAM2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91093128; PubMed=1989895;			
RA	Kallunki T., Ikonen J., Chow L.T., Tryggvason K.;			
RT	"Structure of the human laminin B2 chain gene reveals			
RT	divergence from the laminin B1 chain gene."			
RT	J Biol. Chem. 266:221-228(1991).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=86198245; PubMed=3360804;			
RA	Pikkarainen T., Kallunki T., Tryggvason K.;			
RT	"Human laminin B2 chain. Comparison of the complete amino acid			
RT	sequence with the B1 chain reveals variability in sequence homology			
RT	between different structural domains."			
RT	J Biol. Chem. 263:6751-6758(1988).			
RL	[3]			
RN	SEQUENCE OF 1393-1609 FROM N.A.			
RP	MEDLINE=89169663; PubMed=3234037;			
RA	Fukushima Y., Pikkarainen T., Kallunki T., Eddy R.L., Byers M.G.;			
RT	Haley L.V., Henry W.M., Tryggvason K., Shows T.B.;			
RT	"Isolation of a human laminin B2 (LAM2) cDNA clone and assignment of			
RT	the gene to chromosome region 1q25--q31."			
RT	Cytogenet. Cell Genet. 48:137-141(1988).			
RL	[4]			
RN	SEQUENCE OF 1282-1609 FROM N.A.			
RP	TISSUE=Endothelial cells;			
RA	MEDLINE=92216129; PubMed=1806043;			
RT	Santos C.U.S., Sabbaga J., Brentani R.;			
RT	"Differences in human laminin B2 sequences."			
RT	DNA Seq. 1:275-277(1991).			
CC	-1- FUNCTION: Binding to cells via a high affinity receptor, laminin			
CC	is thought to mediate the attachment, migration, and organization			
CC	of cells into tissues during embryonic development by interacting			
CC	with other extracellular matrix components.			
CC	-1- SUPPLEMENT: Laminin is a complex glycoprotein, consisting of three			
CC	different polypeptide chains (alpha, beta, gamma), which are bound			
CC	to each other by disulfide bonds into a cross-shaped molecule			
CC	comprising one long and three short arms with globules at each			
CC	end.			
CC	THE GAMMA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ),			
CC	LAMININ-2 (MEROSIN), LAMININ-3 (S-LAMININ), LAMININ-4 (S-MEROSIN),			
CC	LAMININ-6 (K-LAMININ) AND LAMININ-7 (KS-LAMININ).			
CC	-1- SUBCELLULAR LOCATION: Extracellular.			

```

CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
CC COMPONENT).
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -1- DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -1- SIMILARITY: CONTAINS 1 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC or send an email to license@lsb-sib.ch.
CC -----
DR EMBL: M55210; AAA59492.1; JOINED.
DR EMBL: M55217; AAA59492.1; JOINED.
DR EMBL: M55201; AAA59492.1; JOINED.
DR EMBL: M55211; AAA59492.1; JOINED.
DR EMBL: M55212; AAA59492.1; JOINED.
DR EMBL: M55213; AAA59492.1; JOINED.
DR EMBL: M55214; AAA59492.1; JOINED.
DR EMBL: M55215; AAA59492.1; JOINED.
DR EMBL: M55216; AAA59492.1; JOINED.
DR EMBL: M55192; AAA59492.1; JOINED.
DR EMBL: M55193; AAA59492.1; JOINED.
DR EMBL: M55194; AAA59492.1; JOINED.
DR EMBL: M55195; AAA59492.1; JOINED.
DR EMBL: M55196; AAA59492.1; JOINED.
DR EMBL: M55197; AAA59492.1; JOINED.
DR EMBL: M55198; AAA59492.1; JOINED.
DR EMBL: M55200; AAA59492.1; JOINED.
DR EMBL: M55202; AAA59492.1; JOINED.
DR EMBL: M55203; AAA59492.1; JOINED.
DR EMBL: M55205; AAA59492.1; JOINED.
DR EMBL: M55206; AAA59492.1; JOINED.
DR EMBL: M55207; AAA59492.1; JOINED.
DR EMBL: M55208; AAA59492.1; JOINED.
DR EMBL: M55209; AAA59492.1; JOINED.
DR EMBL: J03202; AAA59492.1; JOINED.
DR EMBL: M27654; AAA59489.1; JOINED.
DR EMBL: X13939; CAA32122.1; JOINED.
DR PIR: S13548; MHUB2.
DR HSSP: P02468; ILE.
DR GeneW: HGNC:6492; LAMC1.
DR MIM: 150290.
DR InterPro: IPR004089; Chmtaxis_transd.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001886; LAMNT.
DR InterPro: IPR000034; Laminin_B.
DR Pfam: PF00052; laminin_B_1.
DR Pfam: PF00053; laminin_EGF_10.
DR PRINTS: PR00055; laminin_Nterm; 1.
DR ProDom: PD002082; LAMNT; 1.
DR ProDom: PD003031; Laminin_B; 1.
DR SMART: SM00180; EGF_Lam; 10.
DR SMART: SM00281; LAMNT; 1.
DR SMART: SM00136; LAMNT; 1.
DR PROSITE: PS00022; EGF_1; 8.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01248; LAMININ_TYPE_EGF; 11.
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
FT SIGNAL 1 33
FT CHAIN 34 1609 LAMININ GAMMA-1 CHAIN.
FT DOMAIN 34 285 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 286 341 LAMININ EGF-LIKE 1.

```

```

FT DOMAIN 342 397 LAMININ EGF-Like 2.
FT DOMAIN 398 444 LAMININ EGF-Like 3.
FT DOMAIN 445 494 LAMININ EGF-Like 4.
FT DOMAIN 495 504 LAMININ EGF-Like 5 (N-TERMINAL).
FT DOMAIN 505 689 LAMININ EGF-Like 5 (N-TERMINAL).
FT DOMAIN 690 723 LAMININ EGF-Like 5 (N-TERMINAL).
FT DOMAIN 724 772 LAMININ EGF-Like 6.
FT DOMAIN 773 827 LAMININ EGF-Like 7.
FT DOMAIN 828 883 LAMININ EGF-Like 8.
FT DOMAIN 884 934 LAMININ EGF-Like 9.
FT DOMAIN 935 982 LAMININ EGF-Like 10.
FT DOMAIN 1030 1050 LAMININ EGF-Like 11.
FT DOMAIN 1050 1609 LAMININ II AND I.
FT DOMAIN 1609 1609 COILED COIL (POTENTIAL).
FT DISULFID 286 295 BY SIMILARITY.
FT DISULFID 288 305 BY SIMILARITY.
FT DISULFID 307 316 BY SIMILARITY.
FT DISULFID 319 339 BY SIMILARITY.
FT DISULFID 342 351 BY SIMILARITY.
FT DISULFID 344 367 BY SIMILARITY.
FT DISULFID 370 379 BY SIMILARITY.
FT DISULFID 382 395 BY SIMILARITY.
FT DISULFID 398 410 BY SIMILARITY.
FT DISULFID 400 416 BY SIMILARITY.
FT DISULFID 418 427 BY SIMILARITY.
FT DISULFID 430 442 BY SIMILARITY.
FT DISULFID 445 456 BY SIMILARITY.
FT DISULFID 447 463 BY SIMILARITY.
FT DISULFID 465 474 BY SIMILARITY.
FT DISULFID 477 492 BY SIMILARITY.
FT DISULFID 724 733 BY SIMILARITY.
FT DISULFID 726 740 BY SIMILARITY.
FT DISULFID 742 751 BY SIMILARITY.
FT DISULFID 754 770 BY SIMILARITY.
FT DISULFID 773 781 BY SIMILARITY.
FT DISULFID 775 792 BY SIMILARITY.
FT DISULFID 795 804 BY SIMILARITY.
FT DISULFID 807 825 BY SIMILARITY.
FT DISULFID 828 842 BY SIMILARITY.
FT DISULFID 830 849 BY SIMILARITY.
FT DISULFID 852 861 BY SIMILARITY.
FT DISULFID 864 881 BY SIMILARITY.
FT DISULFID 884 898 BY SIMILARITY.
FT DISULFID 896 905 BY SIMILARITY.
FT DISULFID 907 916 BY SIMILARITY.
FT DISULFID 919 932 BY SIMILARITY.
FT DISULFID 935 947 BY SIMILARITY.
FT DISULFID 947 954 BY SIMILARITY.
FT DISULFID 956 965 BY SIMILARITY.
FT DISULFID 968 980 BY SIMILARITY.
FT DISULFID 983 995 BY SIMILARITY.
FT DISULFID 985 1001 BY SIMILARITY.
FT DISULFID 1003 1012 BY SIMILARITY.
FT DISULFID 1015 1028 BY SIMILARITY.
FT DISULFID 1031 1031 BY SIMILARITY.
FT DISULFID 1034 1034 INTERCHAIN (PROBABLE).
FT DISULFID 1600 1600 INTERCHAIN (PROBABLE).
FT CARBOHYD 60 60 N-LINKED (GLCNAC).
FT CARBOHYD 134 134 N-LINKED (GLCNAC).
FT CARBOHYD 576 576 N-LINKED (GLCNAC).
FT CARBOHYD 650 650 N-LINKED (GLCNAC).
FT CARBOHYD 1022 1022 N-LINKED (GLCNAC).
FT CARBOHYD 1107 1107 N-LINKED (GLCNAC).
FT CARBOHYD 1161 1161 N-LINKED (GLCNAC).
FT CARBOHYD 1175 1175 N-LINKED (GLCNAC).
FT CARBOHYD 1205 1205 N-LINKED (GLCNAC).
FT CARBOHYD 1223 1223 N-LINKED (GLCNAC).
FT CARBOHYD 1241 1241 N-LINKED (GLCNAC).
FT CARBOHYD 1380 1380 N-LINKED (GLCNAC).
FT CARBOHYD 1395 1395 N-LINKED (GLCNAC).
FT CARBOHYD 1439 1439 N-LINKED (GLCNAC).
FT CONFLICT 212 212 F -> I (IN REF. 2).
SQ SEQUENCE 1609 AA: 17606 MM: B098F20FCD97233B CRC64:

```

Query Match 4.6%; Score 251; DB 1; Length 1609;
 Best Local Similarity 19.6%; Pred. No. 4,1e-10;
 Matches 212; Conservative 85; Mismatches 382; Indels 400; Gaps 54;

23 PLRLMLLMAGTAVOVT-----OGTGPDLACKES-----EYHEYTACDSTGSRW 68
 130 PSSINLTLHLKADITVRLKFTSRPESFAIKRREDGPMIPYIYSGSCENTYSKA 169
 69 RYAVPHPPG-----LCTSLPDPVKTGBCSPSCNAGEFLDMKQSCPKCAEGRVSLGTGIR 123
 190 NRGFIKRGDGEQALCTD-----EESDF-----SPLTGVNAFSTP--- 224
 124 FDEWDELPHGFASLSANMELDDSAESTGCTSSKWPREDGYIAPNDECTATIMAVNL 183
 225 -----LEGRPSAYNFNDNSPELDEWYATD-----IRVTL 253
 184 KOSGTVNE-----YYPDSIIFEFYONDOCPNADDSRMKKTTEKGEFHSV 233
 254 NRLNTGDEVENDKVLKSYTALSD-----FVNGCKCKNGHASECKMN-----EFDKL 303
 224 ELNRGNVNL-----YMTTAFSVTKVKKPVLYBNIAITGVAYTSECP 278
 304 VCNCKHNTYGVDCCKLPFNDPMBRATAES-----ASECLPC 342
 279 -----KPGTY-----ADKQSSPCKLCPANSY-----SNKGETSCHQCD 312
 343 DGNRSQCYFDPPELYRSTGHGCTNCQDNTDGAHCERKENEFLGNACSSCH-CS 401
 313 P-----DKYSEKSSSCNWRPA-----CTDKDYFTHT-----ACDANGETQLM 351
 402 PVGSLSTQCDXYR-----CSCKPGVMKDKDRQOPGHSJLEAGCRPCSDPGS1DEC 456
 352 YKMAKPKICSEDLBGAVKLPASGVKTHCPNPGFFFTNNST---COPCYGTSNOSDC 408
 457 NVETGRVCCKNVEG-----FNCERCKPGFFFNLESSNPRGCTPC--FCGHSYVC 504
 409 TRCPAGTEPAVGEYKMMNLTPLTNETTVLSGINPEYKGMGWAVAGHITAGASND 468
 505 TN-----AVGYSV-----YSISTFOIDEDGMRAEDRDSSEASLEMSEROD 546
 469 FMILT-----LVYGFRRPOSVADENREVARITFEVET---LGSVNCLEYMAGVN 518
 547 IAVISDFFPRFTAPRAFLKQVLYSTQN-----LSSFFVDRDRRLSMDLVLSAG 601
 519 SRTNTPVETWKS--KGKOSYIIEENTTSFTW-----AFQ-----RTTEH 559
 602 LRVSPLIAQNSYSEFTTKYVFLHEATDYPMRPAITPPEFOKLNLNLSIKIRGYS 661
 560 EASRKYTDVAKI-----YKIN-----YSIN 576
 662 EESAGYLDVTLASARPGVPATWVESCTCPVYGGOFCMCISGYRRETPNLAGPSPC 721
 577 VTNVANGVASYCRP---CALEASDVGSSTSCAPAGYIDRDSGT---CHSCP--PNTILK 628
 722 VLCAONGHSEPCDETEGVCNCRDNTAGPHCKGSDGYGDSATAGSSQCGPCPGSSC 781
 629 AHQPYGVQ-ACVPCGPGTKNNKIHSLC---YNDCTESNTPTR-----TFNINFSALA 677
 782 AVVPKTEKVVCTNCTGTGTRKC-ELCDDGYGDPJLGRNGEVRLRLCQCSDNIDPNVAVG 840
 678 NVTVLGSGPSTKGLTKYFHH-----FTLSLGNQGRKMSVCDNTDRLIP 724
 841 NENRLTG-----ECLKICTYNAGYCDRCCKDGFGNPLANPADKACACN-----P 888
 725 EESGFSKSTAYVCOAVIIPPEVTGYRAGV-----SSOPYSLAD-RLIGVYTD- 773
 889 YCTMGOSSCNVPGQCECL-PHYVGOCGACDPGYMYLQSGGGERCDHALGSTNQC 947
 774 -----TLDCITSPAEFLHESIGITDVIFFYRSNDVYOSCSGSRSTTRVRCSPKT 825
 948 DIRTGQCECQPGITGQ---HCERC---EVMHFGPGGRCRCD-----CHPR-- 988

QY 826 VPSLL-----POTCS-----DGTDCGCMFHL-----WESAACPLC-----SVADY 864
 DB 989 --GSLSLQCKDRCRCRCRGFPVGNRCDCQCEPNFTYRMSPGQECPCACRYLKVADH 1045
 RESULT 2
 ID LMGL MOUSE STANDARD: PRT: 1607 AA.
 AC P02468;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Laminin gamma-1 chain precursor (Laminin B2 chain).
 GN LAMC1 OR LAMC-1 OR LAMB-2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN RP SEQUENCE FROM N.A.
 RX MEDLINE=88059118; PubMed=3680290;
 RA Sasaki M., Yamada Y.;
 RT "The laminin B2 chain has a multidomain structure homologous to the
 RT B1 chain."
 RL J. Biol. Chem. 262:17111-17117(1987).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89000737; PubMed=3167041;
 RA Durkin M.E., Bartos B.B., Liu S.-H., Phillips S.L., Chung A.E.;
 RT "Primary structure of the mouse laminin B2 chain and comparison with
 RT laminin B1."
 RL Biochemistry 27:5198-5204(1988).
 RN [13]
 RP SEQUENCE OF 1-239 FROM N.A.
 RX MEDLINE=88228071; PubMed=2836421;
 RA Ogawa K., Burdello P.D., Sasaki M., Yamada Y.;
 RT "The laminin B2 chain promoter contains unique repeat sequences and
 RT is active in transient transfection."
 RL J. Biol. Chem. 263:8384-8389(1988).
 RN [14]
 RP SEQUENCE OF 1391-1607 FROM N.A.
 RX MEDLINE=85051302; PubMed=6209134;
 RA Barlow D.P., Green N.M., Kurkinen M., Hogan B.L.M.;
 RT "Sequencing of laminin B chain cDNAs reveals C-terminal regions of
 RT coiled-coil alpha-helix."
 RL EMBO J. 3:2355-2362(1984).
 RN [15]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 771-932.
 RX MEDLINE=96196434; PubMed=8648630;
 RA Steinfeld J., Mayer U., Timpl R., Huber R.;
 RT "Crystal structure of three consecutive laminin-type epidermal growth
 RT factor-like (LE) modules of laminin gamma1 chain harboring the
 RT nidogen binding site."
 RL J. Mol. Biol. 257:644-657(1996).
 RN [16]
 RP STRUCTURE BY NMR OF 824-881.
 RX MEDLINE=96196435; PubMed=8648631;
 RA Baumgartner R., Czisch M., Mayer U., Poeschl E., Huber R.;
 RA Timpl R., Holak T.A.;
 RT "Structure of the nidogen binding LE module of the laminin gamma1
 RT chain in solution."
 RL J. Mol. Biol. 257:658-668(1996).
 CC - FUNCTION: Binding to cells via a high affinity receptor, laminin
 CC is thought to mediate the attachment, migration, and organization
 CC of cells into tissues during embryonic development by interacting
 CC with other extracellular matrix components.
 CC - SUBUNIT: Laminin is a complex glycoprotein, consisting of three
 CC different polypeptide chains (alpha, beta, gamma), which are bound
 CC to each other by disulfide bonds into a cross-shaped molecule
 CC comprising one long and three short arms with globules at each
 CC end.
 CC THE GAMMA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ),
 CC LAMININ-2 (MEROSIN), LAMININ-3 (S-LAMININ), LAMININ-4 (S-MEROSIN),
 CC

CC	LAMININ-6 (K-LAMININ) AND LAMININ-7 (KS-LAMININ).
CC	- SUBCELLULAR LOCATION: Extracellular.
CC	- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR COMPONENT).
CC	- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC	- DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
CC	- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC	- SIMILARITY: CONTAINS 1 LAMININ EGF-LIKE DOMAINS.
CC	- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
CC	-----
DR	EMBL; X05211; CAA8838.1; -. .
DR	EMBL; J03484; AAA39405.1; -. .
DR	EMBL; J02930; AAA39408.1; -. .
DR	EMBL; J03749; AAA39409.1; -. .
DR	PIR; A28469; MMSBZ.
DR	PDB; 1KLO; 20-AUG-97.
DR	PDB; 1TLE; 12-FEB-97.
DR	MGP; MGI:99914; Lamc1.
DR	InterPro: IPRO04089; Chmtaxis_transd.
DR	InterPro: IPRO00561; EGF-like.
DR	InterPro: IPRO01864; LamNT.
DR	InterPro: IPRO00034; LamNT_B.
DR	InterPro: IPRO02049; Laminin_EGF.
DR	Pfam; PF00052; laminin_B; 1.
DR	Pfam; PF00053; laminin_EGF; 10.
DR	Pfam; PF00055; laminin_Nterm; 1.
DR	PRINTS; PR00011, EGFLAMININ.
DR	ProDom; PD002082; LamNT; 1.
DR	ProDom; PD003031; Laminin_B; 1.
DR	SMART; SM00180; EGF_Lam; 9.
DR	SMART; SM00001; EGF_like; 1.
DR	SMART; SM00281; Lamb; 1.
DR	SMART; SM00136; LamNT; 1.
DR	PROSITE; PS00022; EGF_1; 8.
DR	PROSITE; PS01186; EGF_2; 2.
DR	PROSITE; PS01248; LAMININ_TYPE_EGF; 10.
KW	Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil; Laminin EGF-like domain; Cell adhesion; Repeat; Signal; 3D-structure.
FT	SIGNAL 1 33
FT	CHAIN 34 1607 LAMININ GAMMA-1 CHAIN.
FT	DOMAIN 34 283 LAMININ N-TERMINAL (DOMAIN VI).
FT	DOMAIN 284 339 LAMININ EGF-LIKE 1.
FT	DOMAIN 340 395 LAMININ EGF-LIKE 2.
FT	DOMAIN 396 442 LAMININ EGF-LIKE 3.
FT	DOMAIN 443 492 LAMININ EGF-LIKE 4.
FT	DOMAIN 493 502 LAMININ EGF-LIKE 5 (N-TERMINAL).
FT	DOMAIN 503 687 LAMININ DOMAIN IV.
FT	DOMAIN 688 721 LAMININ EGF-LIKE 5 (C-TERMINAL).
FT	DOMAIN 722 770 LAMININ EGF-LIKE 6.
FT	DOMAIN 771 825 LAMININ EGF-LIKE 7.
FT	DOMAIN 826 881 LAMININ EGF-LIKE 8 (NIDOGEN-BINDING).
FT	DOMAIN 882 932 LAMININ EGF-LIKE 9.
FT	DOMAIN 933 980 LAMININ EGF-LIKE 10.
FT	DOMAIN 981 1028 LAMININ EGF-LIKE 11.
FT	DOMAIN 1029 1607 LAMININ II AND I.
FT	DOMAIN 1034 1594 COILED COIL (POTENTIAL).
FT	DISEULFD 340 349 BY SIMILARITY.
FT	DISEULFD 342 365 BY SIMILARITY.
FT	DISEULFD 368 377 BY SIMILARITY.
FT	DISEULFD 380 393 BY SIMILARITY.
FT	DISEULFD 396 408 BY SIMILARITY.
FT	DISEULFD 398 414 BY SIMILARITY.
FT	DISEULFD 416 425 BY SIMILARITY.
FT	DISEULFD 428 440 BY SIMILARITY.

DB	Query Match	Best Local Similarity	Matches	Conservative	Score	DB 1	Length	Indels	Gaps
DB	23 PRMLRLMAGTAFOVT-----OCTGPELHACKS-----EYHYEYTAOCDSTGSRW	4.1%;	19.6%;	83;	225;	DB 1:	1607;	416;	53;
DB	128 PMSIMLTLHLGAKPDIYVRLKFMHSRPSFAIYRTREDGPMIYQYVSGSCENTYSTA	19.6%;	83;	368;	108;	DB 1:	1607;	416;	53;
DB	69 RVAVDPHTG-----LCTSLDPVKGTECSFSCNAEPELDMKQSKPCAEGRYSJGTGIR	19.6%;	83;	368;	108;	DB 1:	1607;	416;	53;
DB	168 NRGFRTRGGDEQALCTD-----EFSDI-----SPLTGVNAVST---	19.6%;	83;	368;	108;	DB 1:	1607;	416;	53;
DB	124 FDEWDELPHGFASLSDAASSTGNCSTSSKWPVRGDIYAETDECTATLMYAVNL	19.6%;	83;	368;	108;	DB 1:	1607;	416;	53;
DB	223 -----LEGRPSAVNFPNSPIVQAGVATD-----IRVTL	19.6%;	83;	368;	108;	DB 1:	1607;	416;	53;

OY 184 KOSGVNFE-----YYPPDSIIFFEFYVNDCCQPNADSKRMKTTEKMEFHSV 233
 DB 252 NRIINTEGDEVNEPKVLSYYIAISD-----FAVGRCKCNHASECVKN-----EFDL 301
 OY 234 ELNRGNVL-----YKRTAFSVMTKVP-----KPLVENIAI 266
 DB 302 MCCKHNTYGVDCCKLPFENDRPWRRAIASASISLPCDCKNGSQCPCYFDPPELYRS--- 358
 OY 267 TGVAVTSECPCKPRTYADKGSFCKLCPANSYSNKGSCSCDCCDDPKSEKSSS--- 323
 DB 359 TG--HGCHCTNCNDNTDGNK-----CERCREN-FFRLGNTRA--CSPCHCSPVSLSTQC 408
 OY 324 -----CNVRA-----CTDKDYFYHT-----ACDANGETOIMYKAPKICSEDL 365
 DB 409 DSYGRCSCKPGVMGDKCRQCPGFHSLTEAGRCPCSCDLRGSTDBCANETGRCYCKKNV 468
 OY 366 GAVKLPAAGVTHCPNCPNPGFCKTNNST---CQPCPYGYSNGSCDCTGCACTEPVAFGE 422
 DB 469 G-----FNCERCKPGFNLSSNPKGCTPC--FCFGHSSVCTN-----AVGYS 509
 OY 423 -YKMNNTLPTMMETTVLSGINFEXKMTGMEVAGDHITTAAGASDNDFMILTLYVYGERP 481
 DB 510 VYDISTFOIDED-----GMRV-----EQRDGESEASLEMSDDQ 543
 OY 482 POSVADT--ENKEVARITFEVETLCSVNCLEF-----MVGVSRTN 522
 DB 544 DIAVISDSYFFRYFLAPVKFLGNOLVSYGONLSFSFVNRDRRLASADLVLEGAGLRVS 603
 OY 523 TPVETWKGK-KGKOSYTYIIEENTTSFTW-----AFQ-----RTTTHASR 563
 DB 604 VPLIAGNSPSETTYKTLFRLHEATDTPMRPALSPFEFOKLNNLNTSIRKIGYSBRTA 663
 OY 564 KYNDVA-----KIYSINVTNV 580
 DB 664 GYIDDTVLOSARPGVPAPVATWESCTCPYVGCGCEFTCLDGRRETPSLGYPVLCIT 723
 OY 581 MNGVASTCRP-----CALIASDVSSCTSCPAGYIIDRDSGT---CHSCP-----P 623
 DB 724 CNGHSETCDPBTGYVCDRCRNTAPNHEKSDGYGDSLTIGTSSDQPCPGSSCAIYP 783
 OY 624 NT--ILKAHOPYGV--QACVPCGPGTRKNNKINSLCNDCTFSRNPTR-----TFNYN 672
 DB 784 KTEKVVCHCTGTAGRCRCELDGDFGDLGS-----NGVRLCRPCQCNMD 833
 OY 673 PSALANTYTLAGPSTSGKLYPHN-----FTLSLGNQGRKMSVCTDNTV 719
 DB 834 PNAVANCNRLTG-----ECLKCIYNAGFYCDRCKEGFNGPLADNADCKKACACN-- 885
 OY 720 DIRIEG--ESGFSKITAVYCAVLIIPRYGYKAGV-----SSQVSLADRLI 767
 DB 886 ----PYGVVOOQSSCNPTGQ--COCL--PMSGRDGTCDPCGYNNQSGQGERCDCHA 937
 OY 768 GVTDTMTDGTISPAELFHLISLIGPDIYFFYRSNDVYQSSGSRITTVRC----- 820
 DB 938 LGSTNOCODIRIGCE-----CQPIQOHCERKETNHFR 973
 OY 821 SPQRTVP-----GSLLL-----PGTCS-----DGTCDGCNPHL-----MESAACP 859
 DB 974 GRECKRCDCDHGSGSLQCKDDGRCRCHREGVGNKRCQCEENFYNNMPCGCECPAC 1032
 RESULT 3
 LMG1_DROME STANDARD: PRT: 1639 AA.
 ID LMG1_DROME
 AC P15215: 024373; 09VTT16;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Laminin gamma-1 chain precursor (Laminin B2 chain).
 GN LAM2 OR LAMC1 OR LAMC1 OR CG3322.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canton-S, and Oregon-R;
 RX MEDLINE=9129161; PubMed=1840513;
 RA Chi H.-C., Jumanaga D., Wang S.-Y., Hul C.-F.;
 RT "Structure of the Drosophila gene for the laminin B2 chain.";
 RL DNA Cell Biol. 10:451-466(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-R;
 RX MEDLINE=89109164; PubMed=2912972;
 RA Chi H.-C., Hul C.-F.;
 RT "Primary structure of the Drosophila laminin B2 chain and comparison
 with human, mouse, and Drosophila laminin B1 and B2 chains.";
 RL J. Biol. Chem. 264:1545-1550(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90037237; PubMed=2808533;
 RA Montell D.J., Goodman C.S.;
 RT "Drosophila laminin: sequence of B2 subunit and expression of all
 three subunits during embryogenesis.";
 RL J. Cell Biol. 109:2441-2453(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amaratunga C., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.H.C., Blazer R.G., Chapple M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Miklos G.L.G.,
 Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkuh L., Beasley E.M.,
 Balow R.M., Basu A., Baxendale J., Bayraktaroglu I., Bolshakov S.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bhatnagar P.,
 Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 Butcher J.M., Butcher S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durkin K.J., Evangelista C.C., Ferrara C., Ferreira S., Fleischmann W.,
 Fouts C., Gabriellian A.E., Garg N.S., Gehlert W.M., Glasser K.,
 Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.I., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.H., Idegram C.,
 Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kelchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp Y., Lin X.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Modarres C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacלב J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Stiden-Kramos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wassarman D.A., Weinstein K.C., Wu D., Yang S., Yao Q.A.,
 Williams S.M., Woodage T., Worley K.C., Xu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [5]
 RP SEQUENCE OF 344-1639 FROM N.A.
 RC STRAIN=Oregon-R;
 RX MEDLINE=88303364; PubMed=3405777;
 RA Chi H.-C., Hul C.-F.;
 RT "cDNA and amino acid sequences of Drosophila laminin B2 chain.";

[illegible]

OY 306 TSCHCDDPKYSEKSSSCVNPACTDKDYFTHHACADGETOLMYKMAKICSEDL 365
 DB 456 HGC000CGD---SGSHQWT-PAC-----DTEGICF-----CKENVE 489
 OY 366 GAVKLPASGVKTHCPNCPGFF---KTNNSTCQPCPYGSGNSGSDCTRCPCAGTEPAV--- 419
 DB 490 G-----RRCNECKPGFENLDKNNRRCGTPC--FCGHTSECMTPAGSYSVTSN 537
 OY 420 -GFEYKMMNTLPNTMETTVSGINFEYKMGTEWVAG--DHITYAAGASDNDMILLTV 476
 DB 538 FNKEFERMTAADLNOREVDIKN---QXRSIGTTAGNHEVY----- 577
 OY 477 PGFRPPOSVMAD---TENKEVARITFVEPELCSNCLCYMVCNSTRNPTVEFMKGS-- 531
 DB 578 -FQAPDRFLGDORASYNRDKFKLOLVGOVANGVSDVILEGASRISLPIRA-QGNGI 634
 OY 532 --KGASTYITIEENTTSFTWAFQRTFHEASRKYNDYAKI-----YSINVTVM 581
 DB 635 PDGKVEYTERLHEH--HDYOMPS-----QSARGFLSLNLPAIKIRATYSVOGEAIL 687
 OY 582 NGV-----ASYCRPCALASDVSSCTSCPCAGYIDRDSG-----TCH 619
 DB 688 DVELEQTAHGAAGHPATWIEQCTCEGYLGQFCSCAPGRHSRPARCGPMPICPDCH 747
 OY 620 S-----CPPT--ILKAHQPYQACVPCGPGTKNNKIHSLCYNDCTSRNTPRTFNVF 673
 DB 748 GHADIDSEFGRCICQHNTHG--DNCDQCAKFGYNALGG--TPNDC--KRCP----- 794
 OY 674 SALANTVTLAAGPSFTSKGLKYFHHTLSLGNCRKMSVCTDNTDLRIEBS----- 726
 DB 795 -----CENDGACLIQINDPIYICTECPKGFHSGCE 824
 OY 727 --BSGESKSTTANY-----QNAV-----IIPPEVYKAGVSSQPSLADRLIGVTTMTL 775
 DB 825 QCSDFEFGDPTGLIGEVQTCSCDCNGVNDPNAVNCNNTTGECLKIHNTAGEHCDQL 884
 OY 776 DG-ITSPAEFLHESLQIPVIFFYRSNDVYOSCSSGSRSTTRVRCSPQKTPYGSLLPG 834
 DB 885 SGHFGDPLALPH-----GRCRCSCYEAGTQDQSIITRCDQYVGOCCQCPNVIIGRDC--G 938
 OY 835 TCSDG-----TCDGCFHFLMESA 853
 DB 939 ECGGYFNIRISGNCENCCLDPVGSYNSTCDRYSGOCHCRPGVMGRCQCCENYFYGESS 998
 OY 854 AACPLC 859
 DB 999 EGCKPC 1004
 RESULT 4
 TSA4_GIALA STANDARD: PRT: 713 AA.
 ID TSA4_GIALA
 AC P21849, 1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Major surface-labeled trophozoite antigen 417 precursor.
 GN TSA 417.
 OS Giardia lamblia (Giardia intestinalis).
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardia.
 OX NCBI_TaxID=5741;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 30957 / WB;
 RX MEDLINE=90280395; PubMed=2352929;
 RA Gillin F.D., Hagblom P., Harwood J., Alely S.B., Reiner D.S.,
 RA McCaffery M., So M., Guiney D.G.,
 RT "Isolation and expression of the gene for a major surface protein of
 RT Giardia lamblia." Sci. U.S.A. 87:4463-4467(1990).
 RL [2]
 RN
 RP SEQUENCE OF 480-620 FROM N.A.

RC STRAIN=AD-1;
 RX MEDLINE=93314970; PubMed=8325510;
 RA Ey P.L., Mayrhofer G.,
 RT "Two genes encoding homologous 70-kDa surface proteins are present
 RT within individual trophozoites of the binucleate protozoan parasite
 RT Giardia intestinalis." J.
 RL Gene 129:257-262(1993).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ANCHORED TO THE
 CC PLASMA MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE GIARDIA VARIANT SURFACE PROTEIN FAMILY.
 CC -1- DOMAIN: CONTAINS 29 REPEATS OF THE CYXC MOTIF.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.slb.ch).
 CC
 CC EMBL, M33641; AAA02688.1; -;
 CC EMBL, M97488; AAA02581.1; -;
 CC PIR, A35502; A35502.
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR002174; Furin-like.
 CC InterPro: IPR005127; Giardia_VSP.
 CC Pfam: PF03302; VSP; 2.
 CC SMART: SM00181; EGF; 1.
 CC SMART: SM00261; FU; 3.
 CC SMART: SM00261; FU; 3.
 CC Signal: Antigen; Glycoprotein; Transmembrane; Repeat.
 CC CHAIN 1 18 713
 CC
 CC DOMAIN 18 679
 CC TRANSMEM 680 708
 CC DOMAIN 709 713
 CC CARBOHYD 289 289
 CC CARBOHYD 676 676
 CC VARIANT 582 582
 CC VARIANT 606 606
 CC SEQUENCE 713 AA; 72510 MW; 9AD7195843DE5601 CRC64;
 Query Match 3.8%; Score 211.5; DB 1; Length 713;
 Best Local Similarity 21.6%; Pred. No. 1; Le-07;
 Matches 155; Conservative 72; Mismatches 278; Indels 211; Gaps 39;
 OY 4 PGHSHHLSARVGRTERIRPRLMLMAGTAFOVTOGTGPELHACKESBYHYETAC-D 62
 DB 99 PGHSLICSSDDGDGYCTEAPAGYFAPVGAANTDEQSV-----IACGD 138
 OY 63 STGSRMRYAVPHPTGLC-----TSLPDPVKTE-----CSFSCNAGEFLDMKDSQPCGA 112
 DB 139 TTGVITIAAGNTYKGIADCAECASAPDATAGAEKAVATCT-KCGVSKYL-KONVCVKA 195
 OY 113 EGRVSLGTGIRPDEWDELPHGFASLSANMELIDSAESTGCTSSKVPYPGDIYIARTDE 172
 DB 196 Q--CNSGSTNKFVAVDSDENGNKVCSCSDNLNGVA-----NCDTCYSYDESKRTI---K 244
 OY 173 CT-ATLMAVAVNLKSGTYVNFXYPPDSSIIEFFVYNDQCCPN---ADDSRMKMTTEKGM 228
 DB 245 CTKCTDNNVLTCTTSEGT-----SCVADQCKDGGFFPDDSD----- 279
 OY 229 EFHSHVSLNKGNNVLYWRTAFESVMTKVPVLYVRNIAI-----TGYAITSSECFCKPG 281
 DB 280 -----SAGNCKLPCNDSTDS-----IANCACALVSGRSGAALVT-CSACTDG 321
 OY 282 --TYADK---QGSSPCKLCPANSYNNKGETS---CHQCPDXYSEKSSSCVNPACTDKD 334
 DB 322 YKPSADKTCCEAVSNCKTGCACRACSNKGENGVCTDCD-----GSYLLPTSGCID-- 372
 OY 335 YFTYTHACANGPTOLMYKMAKPKICSEDLGAVKLPASGVKT-----HCPPCNPGPFKT 389

23 PRLRLIMAGTAFOYT-----OGTGPRLHCKE-----SEHYEYACDST----- 64
 124 PTVANLILSLGKAYEITVYALKFHTSPRESFAIYKRTYASGHWMEYOYYSASCOCKTGRP 183
 65 -GSHWRAVPHHTGLCHS-LPD--PYKGTGCSFSCNAGEFLDMKDCKPCAEGHYSLGT 120
 184 EGHVLRGEEDERVAECTSESDISPLNGNVAFS-----TLEGHPS--- 224
 121 GIREDEMDELPHGPAASLANMELDASAESTGCTSKWVPRGDYIAFNTDECTATLWYA 180
 225 AYNEFESVYL-----QEWTSID-LILSLDLNLT---FG 254
 181 VNLKSGTYVNEFYPPDSSTIFFEYV-----ONDCCPNA----- 215
 255 DDFKDPRLQSYTY---AVSDPSVGRCCKNGHASECPNAGOLACRCOHNTTGVC 310
 216 -----DSSRMKTKTEK-----GWEHVSVELANGNVILWRTTAFSWMTKVPRVL 260
 311 ERLPLFPQDRPWARGTAEDANECLPCNCSGHSSECTDREL--YRST----- 355
 261 VRIATVAYITSECPCKPGTYADKOGSSPCKLPANYSNKGESCHOCDDPKYSENG 320
 356 -----GHGHCORCR-----DHTGPHCERCEKRYWMSPTKPCOPCD--CHPAG 398
 321 SSSCNVPACTDKDYFYTTACDANGETOAMKAKPKICEDELEGAVKLPAAGVTNCP 380
 399 SLSTIQ-----CDNSG-----VC-----PCKPIVTKMK--CD 422
 381 PCNPFKTNNSCOP-----CPYGSYNSGDCSTRCPAGT----- 415
 423 RCLGFFHSLSEGCRCPCACVAGSLGTCDBRSCNCKENYVGLCDRCRGTENLOPHN 482
 416 -----EPAVGEYKMMNLTPTMETTVLSGINFEXKMGKMGVAGDHI- 458
 483 PVGSSCCCYGHSKYCSPAAGFO-----EHHHS--DFRH-GAGGQJIRSMGVS 528
 459 -----YTAG-----ASDNDFM-----LTLVVP--GFRPP----- 482
 529 KRPLOMSOGLLILRGEBELSAKKFLGDRLSYGGPVLTLQVPPGSPPIQLRLG 588
 483 -----QSVADT--ENKEVARITFEVETLCSVNCLEYNV 515
 589 AGIALSLRPSLPSPQDRPARRVOLLOETSEBESPLPTFHGRILLS----- 639
 516 GVNSTRNTVETWMSKSGKOSTYTIIEENTTSFTMAFORTTHEAS-----BKYN 567
 640 -----NLTAISTVSGCGHGGVLLCEVOLTSAWP--QRELABPASWETCLCPQCTG 693
 568 DVAKIYSINVTVM--NGVASYCRCALASDV-----GSSCTSCPAGY 610
 694 QCFECALGYKREIPHGGRVANCIPCTCNOHGTCDPNTGILCGHHTGSPCERCMGIFY 753
 611 IDRBSG---TCHSCP--PNTILKAHP--YGVQACVPCGREGTKNKHLSICYNDCTFSRNT 665
 754 GNAPSGRADDCQPCPCGCSACATIPESGDVCTHCPGQRRKCES--CED----- 803
 666 TRTYNYSALANTVTLAGSPFTSGIKYFHHFTLSLGN-----QGRMSVCTDNVTL 721
 804 -----GFTDPLGLSGAQ--PCRCQCSGNVDLANAGNCDPHGHCILR--CLYNTGA 853
 722 RIPEGESFESITAYVCOAVIIPPEVTGYKAVSSOPS-----LADRLIGTWTMTLDG 777
 854 HCHCRBGF-----YGSNAVTRPVKCAPCSCDLRGVSSEKTCMP 893
 778 ITSPAEHLHESLGTPDY-----IFEYRSNDVYTOSSGSGSTTIRVCSQKTV 826
 894 VTS-----QCVLPYVSGRDCSCSPGYDLSGCGSCCHPLGSLJENKCHPK--- 943
 827 PGLLLPGTC-----SDTCDCGNFHLMSAACPCLCSADVADNHAIVSCVAGIKTV 880
 944 -----TGCPCRPVGTGOACDRCOLGFGFESIKCRRCRCSPLGAASSQC---HENSTC 994
 881 VWRPKLCSGGISLPQRVYTKTIDFWLKVQ--ISAGTCTAILTLVTLCYFWKKNOKLEY 939

Db 995 VCR-----PGFVGYKCR-----COD-NFPLADGDTGCEQCP-----TCYALVKEBAKL 1038
 QY 940 KYSKLVNATLK--DCDLPAADSCALMEGE 967
 Db 1039 KARLMMEGWLQBRSDCGSPWG-PLDILQGE 1067
 RESULT 6
 PKCS5_MOUSE STANDARD: PRT: 1877 AA.
 ID 004592: 062040: (Rel. 31, Created)
 AC 01-FEB-1995 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 41, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proprotein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)
 DE (Proteinase convertase pc5) (Subtilisin/kexin-like protease pc5)
 DE (convertase pc5) (pc5) (Subtilisin-like propotein convertase 6)
 DE (SPC6).
 GN PKCS5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1] SEQUENCE OF 330-1877 FROM N.A. (ISOFORM PC5B).
 RP STRAIN=ICR; TISSUE=Intestine;
 RC MEDLINE=93327934; PubMed=8335106;
 RA Nakagawa T., Murakami K., Nakayama K.;
 RT "Identification of an isoform with an extremely large Cys-rich region
 of pc5, a kex2-like processing endoprotease.";
 RL FEBS Lett. 327:165-171(1993).
 RN [2] SEQUENCE FROM N.A. (ISOFORM PC5A).
 RP TISSUE=Brain, and Intestine;
 RC MEDLINE=93224489; PubMed=8468318;
 RA Nakagawa T., Hosaka M., Torii S., Watanabe T., Murakami K.,
 RT "Identification and functional expression of a new member of the
 mammalian kex2-like processing endoprotease family: Its striking
 structural similarity to PACED.";
 RL J. Biochem. 113:132-135(1993).
 RN [3] SEQUENCE FROM N.A. (ISOFORM PC5A).
 RP TISSUE=Adrenal cortex;
 RC MEDLINE=93342056; PubMed=8341687;
 RA Lussan J., Vileau D., Hamelin J., Day R., Chretien M., Seidah N.G.;
 RT "cDNA structure of the mouse and rat subtilisin/kexin-like PC5: a
 candidate proprotein convertase expressed in endocrine and
 nonendocrine cells.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6691-6695(1993).
 RN [4] PARTIAL SEQUENCE, AND SUBCELLULAR LOCATION.
 RP MEDLINE=97103178; PubMed=8947550;
 RA De Bie I., Marcinkiewicz M., Malide D., Lazure C., Nakayama K.,
 RT "The isoforms of proprotein convertase pc5 are sorted to different
 subcellular compartments.";
 RL J. Cell Biol. 135:1261-1275(1996).
 RN [5] DEVELOPMENTAL EXPRESSION.
 RP MEDLINE=96293359; PubMed=8698813;
 RA Constam D.B., Calton M., Robertson E.J.;
 RT "Spca, spc6, and the novel protease SPC7 are coexpressed with bone
 morphogenetic proteins at distinct sites during embryogenesis.";
 RL J. Cell Biol. 134:181-191(1996).
 RN [6] DEVELOPMENTAL EXPRESSION.
 RP MEDLINE=97436919; PubMed=9291583;
 RA Rancourt S.L., Rancourt D.E.;
 RT "Murine subtilisin-like proteinase SPC6 is expressed during embryonic
 implantation, somitogenesis, and skeletal formation.";
 RL Dev. Genet. 21:75-81(1997).

Db 1599 REAGVQSSADSIPTN-----VAYIAVAFICVIVWLFEVRCMIAQNSG 1604

Db 1599 REAGVQSSADSIPTN-----VAYIAVAFICVIVWLFEVRCMIAQNSG 1604

0Y 927 -TCYFWKKNQKLEYKYSKLVNATLKDCDLPADSCATWEGED-VEEDT TTSQ.....

```
Db      1647 RLCWAHXYQVPTTRYEK--MN-----DHVNIT GCGPGRNVTG 984
```

QY 985 SNH 987
ZNYNLSQEDFINEDSLSEDEIHSIDS 1693

Db	:	1
1694	TRH	1696

RESULT 8
VS41-GIALA

ID	VS41_GIALA	STANDARD;	PRT;	6887 AA
NC	P92127.			

DT 16-OCT-2001 (Rel. 40, Created)
TT 16-OCT-2001

16-OCT-2001 (Rel. 40, Last sequence update)	40
16-OCT-2001 (Rel. 40, Last annotation update)	40

Variant-specific surface protein VSP41 precursor (CRISP-90)

Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia
NCBI TaxID=5741.

```
[1] _____  
      RN  
      BP SEQUENCE FROM " "  
      _____
```

---COLLECTED FROM N.A.
STRAIN=02-4A1;
MEDI TIME=020017Z

MELOUNE=9/321554; PubMed=9178264;
Papanastasiou P., Bruderer T., Li Y,
Domay: 2

Primary structure and biochemical properties of a variant-specific surface protein of *Giardia* ^a. Koehler P.;

Mol. Biochem. Parasitol. 86:13-27 (1997).
[2]

P CHARACTERIZATION.
MEDLINE=97233006; PubM=

Papanastasiou P., McConville M.J., Ralton J., Koehler R.
"The variant-variant ratio." PubMed=9078242;

glycosylated and palmitoylated protein".

-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
Biochem. J. 322:49-56(1997).

PLASMA MEMBRANE.
-1- PTM: O-GLYCOSYLATED. THE MATCH CIRCUIT -
-1- PLASMA MEMBRANE PROTEIN, ANCHORED TO THE

-1- AT THE REDUCING TERMINUS, AND HEPARAN SULFATE
PTM: PALMITOYLATED GLYCOPROTEIN IS A TRISACCHARIDE WITH GLC

-1- SIMILARITY: BELONGS TO THE GIARDIA VARIANT SURFACE PROTEIN SUBUNIT

-----PROTEIN FAMILY.-----
This SWISS-PROT entry is copyright. It is produced through
between the Swiss Institute for Bioinformatics and the European
Bioinformatics Institute.

the Swiss Institute of Bioinformatics and the European Bioinformatics Institute. There are no use by the

use by non-profit institutions as long as its content is in no way modified and this statement is not removed from the work.

entitles requires a license agreement (See <http://www.isb-sib.ch/announce> or send an email to license@isb-sib.ch).

EMBL: Z83743: CAB06038 1. -

HSSP; 014763; 1D0G.
GlycoSiteDB: P92127.

InterPro: IPR000561; EGF-like.
InterPro: IPR000374

InterPro: IPR005127; Giardia_vsp.
Pfam: PF03200

SMART: SM00181; EGF; 2.

SMART; SM000001; EGF_like; 1.
SMART; SM00261; FU; 3.

Antigen; Repeat; Transmembrane; Glycoprotein; Lipoprotein; Palmitate; Signal.

SIGNAL	1	14	POTENTIAL.
CHAIN	15	697	

DOMAIN	RESIDUE NO.	FUNCTION
TRANSMEM	660-663	EXTRACELLULAR (POTENTIAL)
15	660	VARIANT-SPECIFIC SURFACE PROTEIN VSP4A1.
45	667	

SEQUENCE	POTENTIAL.	CYTOPLASMIC (POTENTIAL).
DOMAIN	681	
682	687	

SEQUENCE 68 / AA; 70857 MW; D892F675D626D7EC CRC64;

every Match	Score	DB 1;	Length
st Local Similarity	193;	687;	
18.48:			

REV. NO. 2.2e-06;

1

[illegible]

FT DISULFID 820 834 BY SIMILARITY.
 FT DISULFID 822 841 BY SIMILARITY.
 FT DISULFID 844 853 BY SIMILARITY.
 FT DISULFID 856 872 BY SIMILARITY.
 FT DISULFID 875 894 BY SIMILARITY.
 FT DISULFID 901 912 BY SIMILARITY.
 FT DISULFID 903 912 BY SIMILARITY.
 FT DISULFID 915 928 BY SIMILARITY.
 FT DISULFID 931 943 BY SIMILARITY.
 FT DISULFID 933 950 BY SIMILARITY.
 FT DISULFID 952 961 BY SIMILARITY.
 FT DISULFID 964 976 BY SIMILARITY.
 FT DISULFID 979 991 BY SIMILARITY.
 FT DISULFID 981 998 BY SIMILARITY.
 FT DISULFID 1000 1009 BY SIMILARITY.
 FT DISULFID 1012 1023 BY SIMILARITY.
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 507 507 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 643 643 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 725 725 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 921 921 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1062 1062 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1168 1168 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1211 1211 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1244 1244 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1321 1321 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1437 1437 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1557 AA; 172723 MW; CAROBS1F8D5EBDZP CRC64;

Query Match 3.54; Score 192; DB 1; Length 1557;
 Best Local Similarity 19.38; Pred. No. 7.4e-06;
 Matches 211; Conservative 99; Mismatches 390; Indels 394; Gaps 54;

QY 20 RIPLMLLLMAGTAQVOTGTPELHACKSEFHEHYEACSTSSRRVAVPHTPGLC 79
 DB 114 QOYPTTTLVLTGKSPDITVRLKISPRESEFTYKTHHDSWEPQF----- 164
 QY 80 TSLDPVAGTECSFCNAGEFLDMKQSCPCAGRYSLGTGIRPDEMDLPHGFASLSA 139
 DB 165 -----YSGSCA-----TYGLS-----DRAPILPGNEATQOC 191
 QY 140 NMEIDSAESTGNCSTSSKWPBGDIATPNTDE-----CTATIMAVALKQSGVNEE- 192
 DB 192 TKRPSDITPTLGTGIAFSTLEGRPSAHAESEVLOKWTASAI-RISLNMNTFGDEVE 250
 QY 193 -----YYPDSIIIEFFV-----ON----- 208
 DB 251 KDOVLRSYY-----AISDFAVGGRCKCNGHASECVSSVDGENRLVCRCEHNTGADLC 306
 QY 209 DDCQPNADSRMKTTEKGEHSEVLEIN-----RGNNVLWRTTAPSVWTKYKPYVL 260
 DB 307 NCELPFYNDRRPRSGT-----SVEANECIACNCSQLSKRCFFDOOLEE----- 350
 QY 261 VANITAVTSECFPCPGTYADKOGSSFCCLPRANSKNGETSCQDPPDYSEK 320
 DB 351 -----TG-HGGHCIDCQGT-----QG-VHCQCCIANMRRRGENYVACG-----CNEIG 394
 QY 321 SSS-----CNVRPACTCK-----DFY-----THACDANGNGEQLMKRAKAKI 359
 DB 395 SLSTQCNCKCKCKCKGVTGRFCDDCLDGFYFSTNGCKNCCEISGSLN-----NDPRC 449
 QY 360 CSEDLBEAVLPLASGVKTHCPNCPGF-----KTNNSTQPCPYGYSYSGSDCTRCP----- 412
 DB 450 DSSGSCSKLNEVG-----RQCDKCKGVPFLSTENPGCTPC-FCFGSSISICNTADGYFA 505
 QY 413 -----AGTEPAVGEYKRWNTL-PTNMETYLISGINEFYGMGMWEVAGD 456
 DB 506 MNSVAFDDKKKMG-QRIGLQDTQWALDKAVASDITNSPYR-----VADE 555
 QY 457 HITTAGASDNDPMLTLVPGFRPPOSV-----MADTENKEVARI-----T 498

DB 556 QFLGDRSSYNODLVFTLLKAKHVTNODVKDIIIVGADNOELSTSTIAQNPPTTEAQT 615
 QY 499 FVEFTLSCVNCLEF-----AVGNSRNTVEVWKSSKQSYTYIIEENTTS 548
 DB 616 YRFR-----VHADPYFGYPRINLEDFIGLS-NITAIKIRGTYSKDIGYLSNVNLGTA 669
 QY 549 -----FTW-----AFORTTPEASRKYNTNVAIYISNV 577
 DB 670 GVAPSANPQATWIEHCECLPFGVQFCESGSEGFRET-----KFGPFNHCICDC 723
 QY 578 TNWANGVASCRCPCALASDVSSCTSPAGYIDROSGT-----CHSP-PN---TILKAH 630
 DB 724 HNSNSCEASGSCICEHNTAGDTCERCARGYGDALQGTEDQKPCPNDDPCILHAD 783
 QY 631 -----QPYGQACVPCGCGTKNNKTHSLCYNDCTFSRNTPTTFNVSALANTYT 681
 DB 784 GDICTECPNGYTGRCRCDCSDGYFGNPKDTECEVCACSGNTDP-----NSIGCDK 836
 QY 682 LAG-----GSPFSSKGL-----KTFHFTLSLGN 706
 DB 837 ITGECKKCIENTHGPNCEKCPGYWGDLIEPRGNQSCGCACTRRRNNDTYLLCNO 896
 QY 707 QGRKMSVCTDWTDLRIPEGESGFSKSYAVQAVIIP-----EV-TGY---RAG 754
 DB 897 QDQGD-CLPNIQIDQDQCAHGFYNTISGLGQCECNDPLGSEGNCTDVTNQGQCKKPG 955
 QY 755 VSSQPSVLADRLIGVTTMDLTGTSRPAELFHEISIGIPDVIFFRSNDVQSC--SSGR 812
 DB 956 VTGQ-----RCDRCADYHFGFSANQ-CQPCDEYIGS-----ENQOCDVNSG- 996
 QY 813 STIRVRQSPQKTPVPSLLPCTGSDGTCDCG--NFHFLMSAAACPLCSVADYHAIVSS 870
 DB 997 -----QCLCKENVEGR-----RCDQCAENRIGITQGLPC-----DD 1028
 QY 871 CVAGIQKTYVWRE 884
 DB 1029 CYLLIOSRVNVFRE 1042

RESULT 10
 LM33_HUMAN
 ID LM33_HUMAN STANDARD; PRT: 1587 AA.
 AC Q9Y6N6;
 DT 15-JUN-2002 (Ref. 41, Created)
 DT 15-JUN-2002 (Ref. 41, Last sequence update)
 DT 15-JUN-2002 (Ref. 41, Last annotation update)
 DE Laminin gamma-3 chain precursor (Laminin 12 gamma 3).
 GN LM33.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Koch M., Olson P.F., Albus A., Jin W., Hunter D.D., Brunken W.J.,
 RA Bureson R.E., Champilaud M.F.;
 RT "Characterization and expression of the laminin gamma3 chain: a novel,
 RT non-basement membrane-associated, laminin chain.";
 RT J. Cell Biol. 145:605-618(1999).
 CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
 CC is thought to mediate the attachment, migration, and organization
 CC of cells into tissues during embryonic development by interacting
 CC with other extracellular matrix components.
 CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
 CC different polypeptide chains (alpha, beta, gamma), which are bound
 CC to each other by disulfide bonds into a cross-shaped molecule
 CC comprising one long and three short arms with globules at each
 CC end.
 CC THE GAMMA-3 CHAIN IS A SUBUNIT OF LAMININ-12.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: Broadly expressed in: skin, heart, lung, and

the reproductive tracts.
 CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.

CC -1- DOMAIN: DOMAIN IV IS GLOBULAR.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).

CC -1- SIMILARITY: CONTAINS 1 LAMININ EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AF041835; AAD36991.1; .
 CC HSSP: P02468; ITLE.
 CC Gene: HGNC:6494; LAMC3.

DR MIM: 604349; .
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001886; LAMNT.
 DR InterPro: IPR00034; Laminin_B.
 DR InterPro: IPR002049; Laminin_EGF.

DR Pfam: PF00052; Laminin_B; 1.
 DR Pfam: PF00055; Laminin_EGF; 10.
 DR Pfam: PF00055; Laminin_Nterm; 1.
 DR PRINTS: PR0011; EGF_LAMININ.

DR PRODOM: PD002082; LAMNT; 1.
 DR PRODOM: PD003031; Laminin_B; 1.
 DR SMART: SM00180; EGF_Lam; 10.
 DR SMART: SM00181; EGF_Like; 1.
 DR SMART: SM00281; LAMNT; 1.
 DR SMART: SM00336; LAMNT; 1.

DR PROSITE: PS00022; EGF_1; 7.
 DR PROSITE: PS01248; LAMININ; EGF_2; 2.
 DR PROSITE: PS01248; LAMININ; TYPE EGF; 10.
 DR PROSITE: PS01248; LAMININ; Basement membrane; Extracellular matrix; Coiled coil;
 DR PROSITE: PS01248; LAMININ; Cell adhesion; Repeat; Signal.

KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
 KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

FT SIGNAL 1 19
 FT CHAIN 20 1587
 FT DOMAIN 20 270
 FT DOMAIN 271 326
 FT DOMAIN 327 382
 FT DOMAIN 383 429
 FT DOMAIN 430 479
 FT DOMAIN 480 489
 FT DOMAIN 490 672
 FT DOMAIN 673 706
 FT DOMAIN 707 754
 FT DOMAIN 755 809
 FT DOMAIN 810 865
 FT DOMAIN 866 916
 FT DOMAIN 917 964
 FT DOMAIN 965 1013
 FT DOMAIN 1014 1587
 FT DOMAIN 1071 1141
 FT DOMAIN 1200 1229
 FT DOMAIN 1424 1504
 FT DOMAIN 1535 1579
 FT DOMAIN 1587 1587

Best Local Similarity 19.0%; Pred. No. 7.6e-06; Indels 404; Gaps 65;
 Matches 204; Conservative 114; Mismatches 354; Indels 404; Gaps 65;

32 AGTAFOVTOGIGPELHACKSEHYEYACDSTGSRKVAVPHITGLCTSLDPPYKREC 91
 70 AGACRCRCADAPORH--HNSAYLTDHSDQ--STW-- --OSPMAGVYIP--TSV 118
 92 SPSCNAGE-----FLDMKDSKPCABEGRYSLGIGIFED--EMDELPHGRASLANKELD 145
 119 NITLRGKAVETIYVRLKPHTRSPEFAIYRS--RADPWE--PYGFYSAS----- 166
 146 SAABSTGNTSSKMWVRGDIYAFNDE--CTATLWAVNAKOSGVNNE-----Y 193
 167 -----CKTIGREGQYLARGEDERVAFTSTFFS--DISPLSGVAVFSTLEGRSAV 217
 194 YPDSIIFEFYVNDQCCPNADDSRMKMTTEKGEFHSVELNR--GNVL----- 242
 218 NFESPGLOE-----WVSTEL--LISDLRNFEGDIDFEDPKVLO 256
 243 --YKRTAFSWYKVRKPLVRLVNAITGVAYTSECPKPGTYADKOSSPCKLCPANSY 300
 257 SYVAVDSFVSGKCKCN-----GHASECPD-----DVAGQLACR--COHNTT 297
 301 SNKGETSCHQDP--DKYSEKSSS-----CNY--RPACT--DKDYFTT----- 338
 298 G-----TDCERCLPFQDPKPMAGTAENAHCLPCNCSGRSEECTEDRELFRSTGHGRCH 353
 339 -----HTA-----CDANGETQIMYKMAKPKI-----CEDLEGAVKLPASGVKT 377
 354 HCRDHTAGHRCQEN-----FYHM--DPRMPCQCCQASGLIHQCDDTGFCACKPTV 407
 378 --HCPKPCNGFFKTNNSICOPC--PYGSYS-----NSDCTRCPCAGTE 416
 408 TGKCDRCCLPFQSHLSSEGCRCCTCPNAGSLDTCPRSGRCPCKENYEGMLCRCPGT 466
 417 PAVGEFKKMNLTPTN-----METVLSGINE-----YKGMTG--EVA 454
 467 -----FNLPNHPAGSSCFCHGSKVCASIAQOVNHILSDPHQACAEGMWARSV 517
 455 G-----DHITTAAG--ASDNDP-----MITLVYRGPFRPSVADT 489
 518 GSEHSPOKSPNGVLLSEDEBELTAPKFLGDRFSGYQPLLTFRVPPGSDPLPVOLRL 577
 490 ENKEVARTFVEETLGSVNCLEIFPMGVNSRNTPEVETKSGSKOSYTYIIEENTTS 548
 578 EGTGLA-----LSLRHSSLSGQDARASGGGAQVLOETSEDEVAP 618
 549 --FTWAFQRTTHESRKY-----TNDVAKIY--SINTVNMNV--ASYCRPALBAS 596
 619 PLPFFHRLANLTLSLKURSPGSPAGPVLEVALTSARGLSPRASWEICSCPTG 678
 597 DVGSSTSCSPAGYIDRDSG-----TCH--SCPPNT--ILKAPRYGVQACVPCG 642
 679 YTGQFSCAPGYKRMPOGGPYASCVPCTCNQGTCDPNTGICVCSHNIEG--PSCERCL 737
 643 PGRKNKTHSLCYNDCTFSRNTPTFTFYNFSALANTVYLAGRSTFSKGLKFFHFTLS 702
 738 PGFYGNPFAQA--DDC--QPCPCPG-----QSACTTIP

DB 922 LGSEDOCHPRTGCTCRPGVTGACDRCLGFGSSINGCRACRCSPLGASAC 977

RESULT 11
LMB1_HUMAN STANDARD: PRT: 1786 AA.
AC P07942:
DT 01-AUG-1988 (Rel. 08, Created)
DT 15-JUN-2002 (Rel. 08, Last sequence update)
DE Laminin beta-1 chain precursor (Laminin BI chain).
CN LMB1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE FROM N.A.
RX MEDLINE-90368768; PubMed-1975589;
RT "Structure of the human laminin BI chain gene."
RL J. Biol. Chem. 265:15611-15616(1990).
RN [2]
RP MEDLINE FROM N.A.
RX MEDLINE-87280097; PubMed-3611077;
RA Plikarainen T., Eddy R., Fukushima Y., Byers M., Shows T.,
RT "Human laminin BI chain. A multidomain protein with gene (LMB1)
J. Biol. Chem. 262:10454-10462(1987).
RN [3]
RP MEDLINE-1276-1709 FROM N.A.
RX MEDLINE-88021029; PubMed-3661559;
RA Jaye M., Modi W.S., Rieck G.A., Mudd R., Chiu I.M., O'Brien S.J.,
RT "Isolation of a cDNA clone for the human laminin-BI chain and its
gene localization."
RL Am. J. Hum. Genet. 41:605-615(1987).
CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
is thought to mediate the attachment, migration, and organization
of cells into tissues during embryonic development by interacting
with other extracellular matrix components.
CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
different polypeptide chains (alpha, beta, gamma), which are bound
to each other by disulfide bonds into a cross-shaped molecule
comprising one long and three short arms with globules at each
end.
CC THE BETA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ), LAMININ-
2 (MEROSIN), AND LAMININ-6 (K-LAMININ).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
COMPONENT).
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -1- SIMILARITY: CONTAINS VI AND IV ARE GLOBULAR.
CC -1- SIMILARITY: CONTAINS 12.5 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -1- SIMILARITY: CONTAINS 1 LAMININ EGF-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
DR EMBL: M61951; AAA59486.1; JOINED.
DR EMBL: M58147; AAA59486.1; JOINED.
DR EMBL: M61917; AAA59486.1; JOINED.
DR EMBL: M61918; AAA59486.1; JOINED.
DR EMBL: M61921; AAA59486.1; JOINED.
DR EMBL: M61922; AAA59486.1; JOINED.

DR EMBL: M61923; AAA59486.1; JOINED.
DR EMBL: M61924; AAA59486.1; JOINED.
DR EMBL: M61925; AAA59486.1; JOINED.
DR EMBL: M61926; AAA59486.1; JOINED.
DR EMBL: M61927; AAA59486.1; JOINED.
DR EMBL: M61928; AAA59486.1; JOINED.
DR EMBL: M61929; AAA59486.1; JOINED.
DR EMBL: M61930; AAA59486.1; JOINED.
DR EMBL: M61931; AAA59486.1; JOINED.
DR EMBL: M61932; AAA59486.1; JOINED.
DR EMBL: M61933; AAA59486.1; JOINED.
DR EMBL: M61934; AAA59486.1; JOINED.
DR EMBL: M61935; AAA59486.1; JOINED.
DR EMBL: M61936; AAA59486.1; JOINED.
DR EMBL: M61937; AAA59486.1; JOINED.
DR EMBL: M61938; AAA59486.1; JOINED.
DR EMBL: M61939; AAA59486.1; JOINED.
DR EMBL: M61940; AAA59486.1; JOINED.
DR EMBL: M61941; AAA59486.1; JOINED.
DR EMBL: M61942; AAA59486.1; JOINED.
DR EMBL: M61943; AAA59486.1; JOINED.
DR EMBL: M61944; AAA59486.1; JOINED.
DR EMBL: M61945; AAA59486.1; JOINED.
DR EMBL: M61946; AAA59486.1; JOINED.
DR EMBL: M61947; AAA59486.1; JOINED.
DR EMBL: M61948; AAA59486.1; JOINED.
DR EMBL: M61949; AAA59486.1; JOINED.
DR EMBL: M61950; AAA59486.1; JOINED.
DR EMBL: M61951; AAA59486.1; JOINED.
DR EMBL: M55378; AAA59485.1; JOINED.
DR EMBL: M55365; AAA59485.1; JOINED.
DR EMBL: M55371; AAA59485.1; JOINED.
DR EMBL: M55372; AAA59485.1; JOINED.
DR EMBL: M55373; AAA59485.1; JOINED.
DR EMBL: M55374; AAA59485.1; JOINED.
DR EMBL: M55375; AAA59485.1; JOINED.
DR EMBL: M55376; AAA59485.1; JOINED.
DR EMBL: M55377; AAA59485.1; JOINED.
DR EMBL: M55378; AAA59485.1; JOINED.
DR EMBL: M55379; AAA59485.1; JOINED.
DR EMBL: M55380; AAA59485.1; JOINED.
DR EMBL: M55381; AAA59485.1; JOINED.
DR EMBL: M55382; AAA59485.1; JOINED.
DR EMBL: M55383; AAA59485.1; JOINED.
DR EMBL: M55384; AAA59485.1; JOINED.
DR EMBL: M55385; AAA59485.1; JOINED.
DR EMBL: M55386; AAA59485.1; JOINED.
DR EMBL: M55387; AAA59485.1; JOINED.
DR EMBL: M55388; AAA59485.1; JOINED.
DR EMBL: M55389; AAA59485.1; JOINED.
DR EMBL: M55390; AAA59485.1; JOINED.
DR EMBL: M55391; AAA59485.1; JOINED.
DR EMBL: M55392; AAA59485.1; JOINED.
DR EMBL: M55393; AAA59485.1; JOINED.
DR EMBL: M55394; AAA59485.1; JOINED.
DR EMBL: M55395; AAA59485.1; JOINED.
DR EMBL: M55396; AAA59485.1; JOINED.
DR EMBL: M55397; AAA59485.1; JOINED.
DR EMBL: M55398; AAA59485.1; JOINED.
DR EMBL: M55399; AAA59485.1; JOINED.
DR EMBL: M55400; AAA59485.1; JOINED.
DR EMBL: M55401; AAA59485.1; JOINED.
DR EMBL: M55402; AAA59485.1; JOINED.
DR EMBL: M55403; AAA59485.1; JOINED.
DR EMBL: M55404; AAA59485.1; JOINED.
DR EMBL: M55405; AAA59485.1; JOINED.
DR EMBL: M55406; AAA59485.1; JOINED.
DR EMBL: M55407; AAA59485.1; JOINED.
DR EMBL: M55408; AAA59485.1; JOINED.
DR EMBL: M55409; AAA59485.1; JOINED.
DR EMBL: M55410; AAA59485.1; JOINED.
DR EMBL: M55411; AAA59485.1; JOINED.
DR EMBL: M55412; AAA59485.1; JOINED.
DR EMBL: M55413; AAA59485.1; JOINED.
DR EMBL: M55414; AAA59485.1; JOINED.
DR EMBL: M55415; AAA59485.1; JOINED.
DR EMBL: M55416; AAA59485.1; JOINED.
DR EMBL: M55417; AAA59485.1; JOINED.
DR EMBL: M55418; AAA59485.1; JOINED.
DR EMBL: M55419; AAA59485.1; JOINED.
DR EMBL: M55420; AAA59485.1; JOINED.
DR EMBL: M55421; AAA59485.1; JOINED.
DR EMBL: M55422; AAA59485.1; JOINED.
DR EMBL: M55423; AAA59485.1; JOINED.
DR EMBL: M55424; AAA59485.1; JOINED.
DR EMBL: M55425; AAA59485.1; JOINED.
DR EMBL: M55426; AAA59485.1; JOINED.
DR EMBL: M55427; AAA59485.1; JOINED.
DR EMBL: M55428; AAA59485.1; JOINED.
DR EMBL: M55429; AAA59485.1; JOINED.
DR EMBL: M55430; AAA59485.1; JOINED.
DR EMBL: M55431; AAA59485.1; JOINED.
DR EMBL: M55432; AAA59485.1; JOINED.
DR EMBL: M55433; AAA59485.1; JOINED.
DR EMBL: M55434; AAA59485.1; JOINED.
DR EMBL: M55435; AAA59485.1; JOINED.
DR EMBL: M55436; AAA59485.1; JOINED.
DR EMBL: M55437; AAA59485.1; JOINED.
DR EMBL: M55438; AAA59485.1; JOINED.
DR EMBL: M55439; AAA59485.1; JOINED.
DR EMBL: M55440; AAA59485.1; JOINED.
DR EMBL: M55441; AAA59485.1; JOINED.
DR EMBL: M55442; AAA59485.1; JOINED.
DR EMBL: M55443; AAA59485.1; JOINED.
DR EMBL: M55444; AAA59485.1; JOINED.
DR EMBL: M55445; AAA59485.1; JOINED.
DR EMBL: M55446; AAA59485.1; JOINED.
DR EMBL: M55447; AAA59485.1; JOINED.
DR EMBL: M55448; AAA59485.1; JOINED.
DR EMBL: M55449; AAA59485.1; JOINED.
DR EMBL: M55450; AAA59485.1; JOINED.
DR EMBL: M55451; AAA59485.1; JOINED.
DR EMBL: M55452; AAA59485.1; JOINED.
DR EMBL: M55453; AAA59485.1; JOINED.
DR EMBL: M55454; AAA59485.1; JOINED.
DR EMBL: M55455; AAA59485.1; JOINED.
DR EMBL: M55456; AAA59485.1; JOINED.
DR EMBL: M55457; AAA59485.1; JOINED.
DR EMBL: M55458; AAA59485.1; JOINED.
DR EMBL: M55459; AAA59485.1; JOINED.
DR EMBL: M55460; AAA59485.1; JOINED.
DR EMBL: M55461; AAA59485.1; JOINED.
DR EMBL: M55462; AAA59485.1; JOINED.
DR EMBL: M55463; AAA59485.1; JOINED.
DR EMBL: M55464; AAA59485.1; JOINED.
DR EMBL: M55465; AAA59485.1; JOINED.
DR EMBL: M55466; AAA59485.1; JOINED.
DR EMBL: M55467; AAA59485.1; JOINED.
DR EMBL: M55468; AAA59485.1; JOINED.
DR EMBL: M55469; AAA59485.1; JOINED.
DR EMBL: M55470; AAA59485.1; JOINED.
DR EMBL: M55471; AAA59485.1; JOINED.
DR EMBL: M55472; AAA59485.1; JOINED.
DR EMBL: M55473; AAA59485.1; JOINED.
DR EMBL: M55474; AAA59485.1; JOINED.
DR EMBL: M55475; AAA59485.1; JOINED.
DR EMBL: M55476; AAA59485.1; JOINED.
DR EMBL: M55477; AAA59485.1; JOINED.
DR EMBL: M55478; AAA59485.1; JOINED.
DR EMBL: M55479; AAA59485.1; JOINED.
DR EMBL: M55480; AAA59485.1; JOINED.
DR EMBL: M55481; AAA59485.1; JOINED.
DR EMBL: M55482; AAA59485.1; JOINED.
DR EMBL: M55483; AAA59485.1; JOINED.
DR EMBL: M55484; AAA59485.1; JOINED.
DR EMBL: M55485; AAA59485.1; JOINED.
DR EMBL: M55486; AAA59485.1; JOINED.
DR EMBL: M55487; AAA59485.1; JOINED.
DR EMBL: M55488; AAA59485.1; JOINED.
DR EMBL: M55489; AAA59485.1; JOINED.
DR EMBL: M55490; AAA59485.1; JOINED.
DR EMBL: M55491; AAA59485.1; JOINED.
DR EMBL: M55492; AAA59485.1; JOINED.
DR EMBL: M55493; AAA59485.1; JOINED.
DR EMBL: M55494; AAA59485.1; JOINED.
DR EMBL: M55495; AAA59485.1; JOINED.
DR EMBL: M55496; AAA59485.1; JOINED.
DR EMBL: M55497; AAA59485.1; JOINED.
DR EMBL: M55498; AAA59485.1; JOINED.
DR EMBL: M55499; AAA59485.1; JOINED.
DR EMBL: M55500; AAA59485.1; JOINED.
DR EMBL: M55501; AAA59485.1; JOINED.
DR EMBL: M55502; AAA59485.1; JOINED.
DR EMBL: M55503; AAA59485.1; JOINED.
DR EMBL: M55504; AAA59485.1; JOINED.
DR EMBL: M55505; AAA59485.1; JOINED.
DR EMBL: M55506; AAA59485.1; JOINED.
DR EMBL: M55507; AAA59485.1; JOINED.
DR EMBL: M55508; AAA59485.1; JOINED.
DR EMBL: M55509; AAA59485.1; JOINED.
DR EMBL: M55510; AAA59485.1; JOINED.
DR EMBL: M55511; AAA59485.1; JOINED.
DR EMBL: M55512; AAA59485.1; JOINED.
DR EMBL: M55513; AAA59485.1; JOINED.
DR EMBL: M55514; AAA59485.1; JOINED.
DR EMBL: M55515; AAA59485.1; JOINED.
DR EMBL: M55516; AAA59485.1; JOINED.
DR EMBL: M55517; AAA59485.1; JOINED.
DR EMBL: M55518; AAA59485.1; JOINED.
DR EMBL: M55519; AAA59485.1; JOINED.
DR EMBL: M55520; AAA59485.1; JOINED.
DR EMBL: M55521; AAA59485.1; JOINED.
DR EMBL: M55522; AAA59485.1; JOINED.
DR EMBL: M55523; AAA59485.1; JOINED.
DR EMBL: M55524; AAA59485.1; JOINED.
DR EMBL: M55525; AAA59485.1; JOINED.
DR EMBL: M55526; AAA59485.1; JOINED.
DR EMBL: M55527; AAA59485.1; JOINED.
DR EMBL: M55528; AAA59485.1; JOINED.
DR EMBL: M55529; AAA59485.1; JOINED.
DR EMBL: M55530; AAA59485.1; JOINED.
DR EMBL: M55531; AAA59485.1; JOINED.
DR EMBL: M55532; AAA59485.1; JOINED.
DR EMBL: M55533; AAA59485.1; JOINED.
DR EMBL: M55534; AAA59485.1; JOINED.
DR EMBL: M55535; AAA59485.1; JOINED.
DR EMBL: M55536; AAA59485.1; JOINED.
DR EMBL: M55537; AAA59485.1; JOINED.
DR EMBL: M55538; AAA59485.1; JOINED.
DR EMBL: M55539; AAA59485.1; JOINED.
DR EMBL: M55540; AAA59485.1; JOINED.
DR EMBL: M55541; AAA59485.1; JOINED.
DR EMBL: M55542; AAA59485.1; JOINED.
DR EMBL: M55543; AAA59485.1; JOINED.
DR EMBL: M55544; AAA59485.1; JOINED.
DR EMBL: M55545; AAA59485.1; JOINED.
DR EMBL: M55546; AAA59485.1; JOINED.
DR EMBL: M55547; AAA59485.1; JOINED.
DR EMBL: M55548; AAA59485.1; JOINED.
DR EMBL: M55549; AAA59485.1; JOINED.
DR EMBL: M55550; AAA59485.1; JOINED.
DR EMBL: M55551; AAA59485.1; JOINED.
DR EMBL: M55552; AAA59485.1; JOINED.
DR EMBL: M55553; AAA59485.1; JOINED.
DR EMBL: M55554; AAA59485.1; JOINED.
DR EMBL: M55555; AAA59485.1; JOINED.
DR EMBL: M55556; AAA59485.1; JOINED.
DR EMBL: M55557; AAA59485.1; JOINED.
DR EMBL: M55558; AAA59485.1; JOINED.
DR EMBL: M55559; AAA59485.1; JOINED.
DR EMBL: M55560; AAA59485.1; JOINED.
DR EMBL: M55561; AAA59485.1; JOINED.
DR EMBL: M55562; AAA59485.1; JOINED.
DR EMBL: M55563; AAA59485.1; JOINED.
DR EMBL: M55564; AAA59485.1; JOINED.
DR EMBL: M55565; AAA59485.1; JOINED.
DR EMBL: M55566; AAA59485.1; JOINED.
DR EMBL: M55567; AAA59485.1; JOINED.
DR EMBL: M55568; AAA59485.1; JOINED.
DR EMBL: M55569; AAA59485.1; JOINED.
DR EMBL: M55570; AAA59485.1; JOINED.
DR EMBL: M55571; AAA59485.1; JOINED.
DR EMBL: M55572; AAA59485.1; JOINED.
DR EMBL: M55573; AAA59485.1; JOINED.
DR EMBL: M55574; AAA59485.1; JOINED.
DR EMBL: M55575; AAA59485.1; JOINED.
DR EMBL: M55576; AAA59485.1; JOINED.
DR EMBL: M55577; AAA59485.1; JOINED.
DR EMBL: M55578; AAA59485.1; JOINED.
DR EMBL: M55579; AAA59485.1; JOINED.
DR EMBL: M55580; AAA59485.1; JOINED.
DR EMBL: M55581; AAA59485.1; JOINED.
DR EMBL: M55582; AAA59485.1; JOINED.
DR EMBL: M55583; AAA59485.1; JOINED.
DR EMBL: M55584; AAA59485.1; JOINED.
DR EMBL: M55585; AAA59485.1; JOINED.
DR EMBL: M55586; AAA59485.1; JOINED.
DR EMBL: M55587; AAA59485.1; JOINED.
DR EMBL: M55588; AAA59485.1; JOINED.
DR EMBL: M55589; AAA59485.1; JOINED.
DR EMBL: M55590; AAA59485.1; JOINED.
DR EMBL: M55591; AAA59485.1; JOINED.
DR EMBL: M55592; AAA59485.1; JOINED.
DR EMBL: M55593; AAA59485.1; JOINED.
DR EMBL: M55594; AAA59485.1; JOINED.
DR EMBL: M55595; AAA59485.1; JOINED.
DR EMBL: M55596; AAA59485.1; JOINED.
DR EMBL: M55597; AAA59485.1; JOINED.
DR EMBL: M55598; AAA59485.1; JOINED.
DR EMBL: M55599; AAA59485.1; JOINED.
DR EMBL: M55600; AAA59485.1; JOINED.
DR EMBL: M55601; AAA59485.1; JOINED.
DR EMBL: M55602; AAA59485.1; JOINED.
DR EMBL: M55603; AAA59485.1; JOINED.
DR EMBL: M55604; AAA59485.1; JOINED.
DR EMBL: M55605; AAA59485.1; JOINED.
DR EMBL: M55606; AAA59485.1; JOINED.
DR EMBL: M55607; AAA59485.1; JOINED.
DR EMBL: M55608; AAA59485.1; JOINED.
DR EMBL: M55609; AAA59485.1; JOINED.
DR EMBL: M55610; AAA59485.1; JOINED.
DR EMBL: M55611; AAA59485.1; JOINED.
DR EMBL: M55612; AAA59485.1; JOINED.
DR EMBL: M55613; AAA59485.1; JOINED.
DR EMBL: M55614; AAA59485.1; JOINED.
DR EMBL: M55615; AAA59485.1; JOINED.
DR EMBL: M55616; AAA59485.1; JOINED.
DR EMBL: M55617; AAA59485.1; JOINED.
DR EMBL: M55618; AAA59485.1; JOINED.
DR EMBL: M55619; AAA59485.1; JOINED.
DR EMBL: M55620; AAA59485.1; JOINED.
DR EMBL: M55621; AAA59485.1; JOINED.
DR EMBL: M55622; AAA59485.1; JOINED.
DR EMBL: M55623; AAA59485.1; JOINED.
DR EMBL: M55624; AAA59485.1; JOINED.
DR EMBL: M55625; AAA59485.1; JOINED.
DR EMBL: M55626; AAA59485.1; JOINED.
DR EMBL: M55627; AAA59485.1; JOINED.
DR EMBL: M55628; AAA59485.1; JOINED.
DR EMBL: M55629; AAA59485.1; JOINED.
DR EMBL: M55630; AAA59485.1; JOINED.
DR EMBL: M55631; AAA59485.1; JOINED.
DR EMBL: M55632; AAA59485.1; JOINED.
DR EMBL: M55633; AAA59485.1; JOINED.
DR EMBL: M55634; AAA59485.1; JOINED.
DR EMBL: M55635; AAA59485.1; JOINED.
DR EMBL: M55636; AAA59485.1; JOINED.
DR EMBL: M55637; AAA59485.1; JOINED.
DR EMBL: M55638; AAA59485.1; JOINED.
DR EMBL: M55639; AAA59485.1; JOINED.
DR EMBL: M55640; AAA59485.1; JOINED.
DR EMBL: M55641; AAA59485.1; JOINED.
DR EMBL: M55642; AAA59485.1; JOINED.
DR EMBL: M55643; AAA59485.1; JOINED.
DR EMBL: M55644; AAA59485.1; JOINED.
DR EMBL: M55645; AAA59485.1; JOINED.
DR EMBL: M55646; AAA59485.1; JOINED.
DR EMBL: M55647; AAA59485.1; JOINED.
DR EMBL: M55648; AAA59485.1; JOINED.
DR EMBL: M55649; AAA59485.1; JOINED.
DR EMBL: M55650; AAA59485.1; JOINED.
DR EMBL: M55651; AAA59485.1; JOINED.
DR EMBL: M55652; AAA59485.1; JOINED.
DR EMBL: M55653; AAA59485.1; JOINED.
DR EMBL: M55654; AAA59485.1; JOINED.
DR EMBL: M55655; AAA59485.1; JOINED.
DR EMBL: M55656; AAA59485.1; JOINED.
DR EMBL: M55657; AAA59485.1; JOINED.
DR EMBL: M55658; AAA59485.1; JOINED.
DR EMBL: M55659; AAA59485.1; JOINED.
DR EMBL: M55660; AAA59485.1; JOINED.
DR EMBL: M55661; AAA59485.1; JOINED.
DR EMBL: M55662; AAA59485.1; JOINED.
DR EMBL: M55663; AAA59485.1; JOINED.
DR EMBL: M55664; AAA59485.1; JOINED.
DR EMBL: M55665; AAA59485.1; JOINED.
DR EMBL: M55666; AAA59485.1; JOINED.
DR EMBL: M55667; AAA59485.1; JOINED.
DR EMBL: M55668; AAA59485.1; JOINED.
DR EMBL: M55669; AAA59485.1; JOINED.
DR EMBL: M55670; AAA59485.1; JOINED.
DR EMBL: M55671; AAA59485.1; JOINED.
DR EMBL: M55672; AAA59485.1; JOINED.
DR EMBL: M55673; AAA59485.1; JOINED.
DR EMBL: M55674; AAA59485.1; JOINED.
DR EMBL: M55675; AAA59485.1; JOINED.
DR EMBL: M55676; AAA59485.1; JOINED.
DR EMBL: M55677; AAA59485.1; JOINED.
DR EMBL: M55678; AAA59485.1; JOINED.
DR EMBL: M55679; AAA59485.1; JOINED.
DR EMBL: M55680; AAA59485.1; JOINED.
DR EMBL: M55681; AAA59485.1; JOINED.
DR EMBL: M55682; AAA59485.1; JOINED.
DR EMBL: M55683; AAA59485.1; JOINED.
DR EMBL: M55684; AAA59485.1; JOINED.
DR EMBL: M55685; AAA59485.1; JOINED.
DR EMBL: M55686; AAA59485.1; JOINED.
DR EMBL: M55687; AAA59485.1; JOINED.
DR EMBL: M55688; AAA59485.1; JOINED.
DR EMBL: M55689; AAA59485.1; JOINED.
DR EMBL: M55690; AAA59485.1; JOINED.
DR EMBL: M55691; AAA59485.1; JOINED.
DR EMBL: M55692; AAA59485.1; JOINED.
DR EMBL: M55693; AAA59485.1; JOINED.
DR EMBL: M55694; AAA59485.1; JOINED.
DR EMBL: M55695; AAA59485.1; JOINED.
DR EMBL: M55696; AAA59485.1; JOINED.
DR EMBL: M55697; AAA59485.1; JOINED.
DR EMBL: M55698; AAA59485.1; JOINED.
DR EMBL: M55699; AAA59485.1; JOINED.
DR EMBL: M55700; AAA59485.1; JOINED.
DR EMBL: M55701; AAA59485.1; JOINED.
DR EMBL: M55702; AAA59485.1; JOINED.
DR EMBL: M55703; AAA59485.1; JOINED.
DR EMBL: M55704; AAA59485.1; JOINED.
DR EMBL: M55705; AAA59485.1; JOINED.
DR EMBL: M55706; AAA59485.1; JOINED.
DR EMBL: M55707; AAA59485.1; JOINED.
DR EMBL: M55708; AAA59485.1; JOINED.
DR EMBL: M55709; AAA59485.1; JOINED.
DR EMBL: M55710; AAA59485.1; JOINED.
DR EMBL: M55711; AAA59485.1; JOINED.
DR EMBL: M55712; AAA59485.1; JOINED.
DR EMBL: M55713; AAA59485.1; JOINED.
DR EMBL: M55714; AAA59485.1; JOINED.
DR EMBL: M55715; AAA59485.1; JOINED.
DR EMBL: M55716; AAA59485.1; JOINED.
DR EMBL: M55717; AAA59485.1; JOINED.
DR EMBL: M55718; AAA59485.1; JOINED.
DR EMBL: M55719; AAA59485.1; JOINED.
DR EMBL: M55720; AAA59485.1; JOINED.
DR EMBL: M55721; AAA59485.1; JOINED.
DR EMBL: M55722; AAA59485.1; JOINED.
DR EMBL: M55723; AAA59485.1; JOINED.
DR EMBL: M55724; AAA59485.1; JOINED.
DR EMBL: M55725; AAA59485.1; JOINED.
DR EMBL: M55726; AAA59485.1; JOINED.
DR EMBL: M55727; AAA59485.1; JOINED.
DR EMBL: M55728; AAA59485.1; JOINED.
DR EMBL: M55729; AAA59485.1; JOINED.
DR EMBL: M55730; AAA59485.1; JOINED.
DR EMBL: M55731; AAA59485.1; JOINED.
DR EMBL: M55732; AAA59485.1; JOINED.
DR EMBL: M55733; AAA59485.1; JOINED.
DR EMBL: M55734; AAA59485.1; JOINED.
DR EMBL: M55735; AAA59485.1; JOINED.
DR EMBL: M55736; AAA59485.1; JOINED.
DR EMBL: M55737; AAA59485.1; JOINED.
DR EMBL: M55738; AAA59485.1; JOINED.
DR EMBL: M55739; AAA59485.1; JOINED.
DR EMBL: M55740; AAA59485.1; JOINED.
DR EMBL: M55741; AAA59485.1; JOINED.
DR EMBL: M55742; AAA59485.1; JOINED.
DR EMBL: M55743; AAA59485.1; JOINED.
DR EMBL: M55744; AAA59485.1; JOINED.
DR EMBL: M55745; AAA59485.1; JOINED.
DR EMBL: M55746; AAA59485.1; JOINED.
DR EMBL: M55747; AAA59485.1; JOINED.
DR EMBL: M55748; AAA59485.1; JOINED.
DR EMBL: M55749; AAA59485.1; JOINED.
DR EMBL: M55750; AAA59485.1; JOINED.
DR EMBL: M55751; AAA59485.1; JOINED.
DR EMBL: M55752; AAA59485.1; JOINED.
DR EMBL: M55753; AAA59485.1; JOINED.
DR EMBL: M55754; AAA59485.1; JOINED.
DR EMBL: M55755; AAA59485.1; JOINED.
DR EMBL: M55756; AAA59485.1; JOINED.
DR EMBL: M55757; AAA59485.1; JOINED.
DR EMBL: M55758; AAA59485.1; JOINED.
DR EMBL: M55759; AAA59485.1; JOINED.
DR EMBL: M55760; AAA59485.1; JOINED.
DR EMBL: M557

SMART; SM00180; EGF Lam; 11.
 DR SMART; SM00180; EGF Lam; 11.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01248; EGF_2; 2.
 DR PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
 DR Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
 KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
 FT CHAIN 1 21 LAMININ BETA-1 CHAIN.
 FT DOMAIN 22 1786 LAMININ N-TERMINAL (DOMAIN VI).
 FT DOMAIN 22 270 LAMININ EGF-LIKE 1.
 FT DOMAIN 271 334 LAMININ EGF-LIKE 2.
 FT DOMAIN 335 397 LAMININ EGF-LIKE 3.
 FT DOMAIN 398 457 LAMININ EGF-LIKE 4.
 FT DOMAIN 458 509 LAMININ EGF-LIKE 5 (INCOMPLETE).
 FT DOMAIN 510 540 LAMININ EGF-LIKE 6.
 FT DOMAIN 541 771 LAMININ EGF-LIKE 7.
 FT DOMAIN 773 820 LAMININ EGF-LIKE 8.
 FT DOMAIN 821 866 LAMININ EGF-LIKE 9.
 FT DOMAIN 867 916 LAMININ EGF-LIKE 10.
 FT DOMAIN 917 975 LAMININ EGF-LIKE 11.
 FT DOMAIN 1028 1083 LAMININ EGF-LIKE 12.
 FT DOMAIN 1084 1131 LAMININ EGF-LIKE 13.
 FT DOMAIN 1132 1178 LAMININ I.
 FT DOMAIN 1179 1397 LAMININ II.
 FT DOMAIN 1398 1430 LAMININ ALPHA.
 FT DOMAIN 1431 1786 COILED COIL (POTENTIAL).
 FT DOMAIN 1216 1315 COILED COIL (POTENTIAL).
 FT DOMAIN 1353 1781 COILED COIL (POTENTIAL).
 FT DOMAIN 1442 1781 BY SIMILARITY.
 FT DISULFID 271 280 BY SIMILARITY.
 FT DISULFID 273 298 BY SIMILARITY.
 FT DISULFID 300 309 BY SIMILARITY.
 FT DISULFID 312 332 BY SIMILARITY.
 FT DISULFID 335 344 BY SIMILARITY.
 FT DISULFID 337 362 BY SIMILARITY.
 FT DISULFID 365 374 BY SIMILARITY.
 FT DISULFID 377 395 BY SIMILARITY.
 FT DISULFID 398 411 BY SIMILARITY.
 FT DISULFID 400 426 BY SIMILARITY.
 FT DISULFID 428 437 BY SIMILARITY.
 FT DISULFID 440 455 BY SIMILARITY.
 FT DISULFID 458 472 BY SIMILARITY.
 FT DISULFID 460 479 BY SIMILARITY.
 FT DISULFID 481 490 BY SIMILARITY.
 FT DISULFID 493 507 BY SIMILARITY.
 FT DISULFID 773 785 BY SIMILARITY.
 FT DISULFID 794 803 BY SIMILARITY.
 FT DISULFID 806 818 BY SIMILARITY.
 FT DISULFID 821 833 BY SIMILARITY.
 FT DISULFID 833 840 BY SIMILARITY.
 FT DISULFID 842 851 BY SIMILARITY.
 FT DISULFID 854 864 BY SIMILARITY.
 FT DISULFID 867 876 BY SIMILARITY.
 FT DISULFID 883 895 BY SIMILARITY.
 FT DISULFID 886 895 BY SIMILARITY.
 FT DISULFID 898 914 BY SIMILARITY.
 FT DISULFID 933 933 BY SIMILARITY.

Query Match 3.2%; Score 174; DB 1; Length 1786;
 Best Local Similarity 20.9%; Pred. No. 0.00018; Indels 290; Gaps 46;
 Matches 168; Conservative 69; Mismatches 278;

OY 275 CPCKPQY----ADKQSSPCKLCPANSSYNGKSGCHQCDPKYSE-----KGSSSC 324
 DB 374 CECKPQYQHPRDIRPNCFCRKCCTCDPAGSNEGIC-----DSTDFSTGLIAGQCRC 428
 OY 325 --NVRPACTD--KDYFY-----THACDANG-----ETOLMY----- 352
 DB 429 KLVNBEHCDVCKEGFYDLSSEDPGCKSCACNPLGTPGANGCDSTGHCYCKRLVTCG 488
 OY 353 -----KWAPKICSEDLBGAVKLPASGVKTHCPC-----NPGFFTKNNSTCO 395

DB 489 HCDQCLPEHNG-----LSNDLDG-----CRPDGDLGALNNSCFAESGQSCSR 532
 OY 396 PCPYGYSN----- 592
 DB 533 PHMIGROCNVEEGYFATLDHLYAEAEANLPGVSIYEROYIODRLPMTGAGFYRV 592
 OY 413 AGTPAVGFEYKMMNLTLPMTETVLSGINFEEKMTGWEVAGDHIYTAAGASNDNFIL 472
 DB 593 EGAV-----LIEF-FIDNIPYSMEYDIL--LTFEPQLPDHWEK-----VI 629
 OY 473 TLVVDGFRPPOS-----VADPENKEV-----ARITVFETLGSVNCLEXFVGVN--SRT 521
 DB 630 TVORGRIPSTSRGNTIPDDDNVSLSPGSRVYLPRVC-----FEGGTNYTRL 682
 OY 522 NTPVETWKSQKOSYIIE-----ENTTSFTW-AFORITFHFA 561
 DB 683 ELPOYTSSDSDVESPYTLIDSLVMPYCKSLDITVVGSGDGVVTSAMETFORRYCLEN 742
 OY 562 SRKYT-----NDVAK--IYSINVTNMGVASYCRP-----CALASDVSS 601
 DB 743 SRSVVKTPMTDVCNRIITPISALLHQTLGACEDPQGLSSVCDPNGGOCRCRPNVGR 802
 OY 602 CTSCPAGYIYDRSG-----TCH-----SCPPTILKAHOPYVOA-----CYP----- 640
 DB 803 CNRCAPCTGFGFGSGCKPCECHLQGSVNAFCNPVT--GQCHFGQVYAROCDCRLPHWGF 861
 OY 641 --CGPGRKNKIKHSLYCNDCTFSRNTPTFTYVNSALANVTLAGGSPFSKLGKTFH 698
 DB 862 PSCPPQCCNG--HA--DDC-----DPVT-----GRCLMCQDYTMGN 894
 OY 699 FTLSLGNQCKKMSVCTDNTDLRIPEG-ESG--FESKSI-----TAYVQAVIIPPE 747
 DB 895 CERCIAGYGGPILIGSDHCRPCPCPDGPRGQFARSCYODPVTLOLACVCDPCTISR 954
 OY 748 ----VTGY-----KAGVSSOPVSLADRLIGVTTDMLDITGTPAE--LFLHLSLGIPIVI 796
 DB 955 CDDCASGYFGNPSVEGSCQPCQCHNN-IDTTPRACDKETGCKLCLVHTGEGHCQFCR 1013
 OY 797 FFYRSNDYTQSSS-----GRSTTVRCPSP-----QKTVPSLLPCTGSDGCDGCGNFH 848
 DB 1014 FGYYGALHODCRKVCNLTGVQEHCHGSDQCDKATGOCICLPVNYGO--NCDRCAPN- 1071
 OY 849 LME--SAAACPLCSVADYHAIYSSC 871
 DB 1072 TWQLASGTGDCPCNCMAHSPSC 1096

RESULT 12
 LMA2_MOUSE STANDARD; PRT: 3106 AA.
 ID LMA2_MOUSE
 AC 060675; Q05003; Q64061;
 DT 01-NOV-1997 (Rel. 35, last sequence update)
 DT 01-NOV-1997 (Rel. 35, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Laminin alpha-2 chain precursor (Laminin M chain) (Merosin heavy chain).
 GN LMA2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OC NCBI_Taxid=10090;
 OX RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Embryo, and Heart;
 MEDLINE=95316259; PubMed=7795883;
 RA Yamada Y.; Utani A.; Sugiyama S.; Doi T.; Polistina C.;
 RT "Cloning and expression of laminin alpha 2 chain (M-chain) in the mouse."
 RL Matrix Biol. 14:447-455(1995).
 RN 121
 RP SEQUENCE OF 2162-2279 FROM N.A.


```

FT DISULFID 965 961 BY SIMILARITY.
FT DISULFID 983 992 BY SIMILARITY.
FT DISULFID 995 1007 BY SIMILARITY.
FT DISULFID 1010 1019 BY SIMILARITY.
FT DISULFID 1012 1026 BY SIMILARITY.
FT DISULFID 1028 1037 BY SIMILARITY.
FT DISULFID 1040 1053 BY SIMILARITY.
FT DISULFID 1056 1068 BY SIMILARITY.
FT DISULFID 1058 1075 BY SIMILARITY.
FT DISULFID 1077 1086 BY SIMILARITY.
FT DISULFID 1089 1099 BY SIMILARITY.
FT DISULFID 1416 1425 BY SIMILARITY.
FT DISULFID 1418 1432 BY SIMILARITY.
FT DISULFID 1435 1444 BY SIMILARITY.
FT DISULFID 1447 1462 BY SIMILARITY.
FT DISULFID 1465 1480 BY SIMILARITY.
FT DISULFID 1493 1490 BY SIMILARITY.
FT DISULFID 1493 1502 BY SIMILARITY.
FT DISULFID 1505 1520 BY SIMILARITY.
FT DISULFID 1523 1535 BY SIMILARITY.
FT DISULFID 1525 1542 BY SIMILARITY.
FT DISULFID 1544 1553 BY SIMILARITY.
FT DISULFID 1556 1567 BY SIMILARITY.
FT DISULFID 1570 1570 INTERCHAIN (PROBABLE).
FT DISULFID 1574 1574 INTERCHAIN (PROBABLE).
FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 359 359 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

Query Match 3.1%; Score 172.5; DB 1; Length 3106;

Best Local Similarity 19.5%; Pred. No. 0.00047; Matches 191; Conservative 80; Mismatches 304; Indels 403; Gaps 57;

```

QY 42 TGPLHACKSEY-----HYEY---ACDSTGRMVAVPHTPG----- 77
DB 836 TGPCEKACAGYRGQPSVPGSGCPQOCNNDLSYSGSDSLGSLCKPCTTGRYCE 895
QY 78 LCTS--LPDPYKGTGEC-----SFS--CN--AGEFLDMKDSCKPCAEGY----- 116
DB 896 LCADYRGDAVNTKNCQPCRCIDJNGSFSECHRTGQC-----ECRPYNGHHCDECKR 949
QY 117 -----SLGTGIRPEWELP---HGFASSLANMELDDSAESTGNCSTSKWVRGDIYAF 168
DB 950 ETEGLQGRG-----LPCNCSFSGSKSD-----CEASGQC---WCOPG--VAGK 990
QY 169 NTDECTATLMYAVNLKSGTGVNEFYYPDSIIFFEFVQNDQCQPNADDSRMKKTTEKGW 228
DB 991 KCDKCA-----HGTFNQ---EGGCIACDCSHLGNNDP-----KT----- 1023
QY 229 EHSVELNRGNVLYWMTAFSVWTKVPRVLVNRNIAITGVAITSECFPCPKPTGVDKOG 288
DB 1024 -----GQCICPPMTG-----EKCECLPMTW---G 1046
QY 289 SSF---KICLPANYS-----NKGETSHOCDPKYSEKSSSON---VRPACTDK 333
DB 1047 HSIYTGKVCVCSVGLASQCNVNTGQCSCH---PKFSGMKCSSECSRGHMVPLCTTLC 1102
QY 334 DYFY---THACDANGETQOLMYKWKAPKICSEDLG--AVKLPASGKVTQCPNCGFF-- 387
DB 1103 DCFLPBGIDATTCDL--ETR-----KCSQSOTGQCSKVNVEG--HCDRCRGRKFG 1150
QY 388 --KTNNSTQCPYGYSG--SDCTRCRPAETEPANGFEYKMMNLTPTNNETVLGGINF 443
DB 1151 LDANKPLGCSGSC---YCFGVTSCSE-----AKGL--IRTWVTL--SDEQITLPLV-- 1194
QY 444 EYKMTGMEVAGDHITAAAGSDNFMILILVYVGPFRPQSVADPNKKNVARI----- 497
DB 1195 -----DEALQHTTKGIA-----FQKPEIVAKMDVVRQELHLEPPYWK 1232
QY 498 --TFVEETLCSVNCCLYFMVGVNSRTNPTVEITWKG--KQKSYTYLIEENTTTSFT 550

```

```

DB 1233 LPQGEKGKIMAYGGKLVAYIFEARDETFGATYVQYIIRGSPETHARITRMAAPLI 1292
QY 551 WAFQRTTH--EASKKYINDAKT-----YSINV-----NVMNCVASYC 588
DB 1293 GQTFHEIEMEKEKYGDDPRISRTYTRDEFDLIDHILIKATYGVNVR--QSR 1350
QY 589 RQCALASDVSGSSCSPAGYIDR-----DSG--TCHSCPNTILKAHPYG-----V 635
DB 1351 SEISMVAVRPHVILAGSPRALITERCDPDPYSGISCTCAGFRLRSEFGRTPGTL 1410
QY 636 QACVPC---GPQTKNNKIHSLCYN-----DQTSRN 663
DB 1411 CTCVPCQCNHSGSQDDPEYVQCNQHHTAGDFCERCALGYGYIRGLPNDQCPACPLI 1470
QY 664 TPTRTFNNFSAANTVILLAGSPSTSGIKYFHFTSLC--GNQGRKMSVCTDNVDL 721
DB 1471 SPNNFS-----PSCVLEGLD--DYCTACPRGEGGYCERCACAGYTG 1512
QY 722 RPEGESGFSKITAVCOAVIIPPEVTGYKAGVSSQPSVLADRLIGVTTMTLDGITSP 781
DB 1513 PSSPQGS-----COECRCDP-----YGLFVP--CDRVYTGICT----- 1543
QY 782 AELFHLIESLGPDIYFYKSNVDVTQSCSSGSRSTIRVCSFQKTYVPSGLLPGTCSGTC 841
DB 1544 -----CRPGAT-----GRKC 1553
QY 842 DGCNHFLEMSAACPIC 859
DB 1554 DCCE--HWHAREGACVFC 1570

```

RESULT 13

```

TS1L_GIALA
ID TS1L_GIALA STANDARD; PRT; 667 AA.
AC Q03185;
DT 01-JUN-1994 (Rel. 29, Created)
DI 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-1998 (Rel. 36, Last annotation update)
DE Major surface trophozoite antigen 11 precursor.
GN TSPII.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinales; Giardia.
OX NCBI_TaxId=5741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate AD-1;
RX MEDLINE=93241215; PubMed=8479449;
RA Ey P.L., Khanna K., Manning P.A., Mayrhofer G.;
RT "A gene encoding a 69-kilodalton major surface protein of Giardia intestinalis trophozoites."
RL Mol. Biochem. Parasitol. 58:247-258(1993).
CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ANCHORED TO THE PLASMA MEMBRANE.
CC -I- DOMAIN: CONTAINS 27 REPEATS OF A CXXC MOTIF.
CC -I- SIMILARITY: BELONGS TO THE GIARDIA VARIANT SURFACE PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL; M95814; AA02687.1; .
CC PIR; A48579; A48579.
CC HSP; P02468; IKLO.
CC InterPro: IPR000561; EGF-like.
CC InterPro: IPR002174; Furin-like.
CC InterPro: IPR005127; Giardia_VSP.
CC Pfam; PF03302; VSP; 1.
CC SMART; SM00181; EGF; 3.

```


FT	DOMAIN	398	457	LAMININ EGF-LIKE 3.
FT	DOMAIN	458	509	LAMININ EGF-LIKE 4.
FT	DOMAIN	510	540	LAMININ EGF-LIKE 5. (INCOMPLETE).
FT	DOMAIN	541	772	LAMININ DOMAIN IV.
FT	DOMAIN	773	820	LAMININ EGF-LIKE 6.
FT	DOMAIN	821	866	LAMININ EGF-LIKE 7.
FT	DOMAIN	867	916	LAMININ EGF-LIKE 8.
FT	DOMAIN	917	975	LAMININ EGF-LIKE 9.
FT	DOMAIN	976	1027	LAMININ EGF-LIKE 10.
FT	DOMAIN	1028	1083	LAMININ EGF-LIKE 11.
FT	DOMAIN	1084	1131	LAMININ EGF-LIKE 12.
FT	DOMAIN	1132	1178	LAMININ EGF-LIKE 13.
FT	DOMAIN	1179	1397	DOMAIN II.
FT	DOMAIN	1398	1430	DOMAIN ALPHA.
FT	DOMAIN	1431	1786	DOMAIN I.
FT	DOMAIN	1216	1315	COILED COIL (POTENTIAL).
FT	DOMAIN	1368	1388	COILED COIL (POTENTIAL).
FT	DOMAIN	1468	1778	COILED COIL (POTENTIAL).
FT	DISULFID	271	280	BY SIMILARITY.
FT	DISULFID	273	298	BY SIMILARITY.
FT	DISULFID	300	309	BY SIMILARITY.
FT	DISULFID	312	332	BY SIMILARITY.
FT	DISULFID	335	344	BY SIMILARITY.
FT	DISULFID	337	362	BY SIMILARITY.
FT	DISULFID	365	374	BY SIMILARITY.
FT	DISULFID	377	395	BY SIMILARITY.
FT	DISULFID	398	411	BY SIMILARITY.
FT	DISULFID	400	426	BY SIMILARITY.
FT	DISULFID	428	437	BY SIMILARITY.
FT	DISULFID	440	455	BY SIMILARITY.
FT	DISULFID	458	472	BY SIMILARITY.
FT	DISULFID	460	479	BY SIMILARITY.
FT	DISULFID	481	490	BY SIMILARITY.
FT	DISULFID	493	507	BY SIMILARITY.
FT	DISULFID	773	785	BY SIMILARITY.
FT	DISULFID	775	792	BY SIMILARITY.
FT	DISULFID	794	803	BY SIMILARITY.
FT	DISULFID	806	818	BY SIMILARITY.
FT	DISULFID	821	833	BY SIMILARITY.
FT	DISULFID	823	840	BY SIMILARITY.
FT	DISULFID	842	851	BY SIMILARITY.
FT	DISULFID	854	864	BY SIMILARITY.
FT	DISULFID	867	876	BY SIMILARITY.
FT	DISULFID	869	883	BY SIMILARITY.
FT	DISULFID	886	895	BY SIMILARITY.
FT	DISULFID	898	914	BY SIMILARITY.
FT	DISULFID	917	933	BY SIMILARITY.
FT	DISULFID	919	944	BY SIMILARITY.
FT	DISULFID	946	955	BY SIMILARITY.
FT	DISULFID	958	973	BY SIMILARITY.
FT	DISULFID	976	990	BY SIMILARITY.
FT	DISULFID	978	997	BY SIMILARITY.
FT	DISULFID	1000	1009	BY SIMILARITY.
FT	DISULFID	1012	1025	BY SIMILARITY.
FT	DISULFID	1084	1096	BY SIMILARITY.
FT	DISULFID	1086	1103	BY SIMILARITY.
FT	DISULFID	1105	1114	BY SIMILARITY.
FT	DISULFID	1117	1129	BY SIMILARITY.
FT	DISULFID	1132	1144	BY SIMILARITY.
FT	DISULFID	1134	1151	BY SIMILARITY.
FT	DISULFID	1153	1162	BY SIMILARITY.
FT	DISULFID	1165	1176	BY SIMILARITY.
FT	DISULFID	1179	1179	BY SIMILARITY.
FT	DISULFID	1182	1182	INTERCHAIN (PROBABLE).
FT	DISULFID	1182	1182	INTERCHAIN (PROBABLE).
FT	DISULFID	1785	1785	INTERCHAIN (PROBABLE).
FT	CARBOHYD	120	120	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	356	356	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	519	519	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	677	677	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1041	1041	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1195	1195	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1279	1279	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1336	1336	N-LINKED (GLCNAC. . .)

FT	CARBOHYD	1343	1343	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1487	1487	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1533	1533	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1542	1542	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1643	1643	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1531	1534	N-LINKED (GLCNAC. . .)
FT	CONFLICT	1749	1749	SGNA -> MEMB (IN REF. 2).
SO	SEQUENCE	1786 AA;	196904 MW;	846671B7B41A474 CRC64;
Query Match				
Best local Similarity 3.1%; Score 170; DR 1; Length 1786;				
Matches 166; Conservative 73; Mismatches 258; Indels 326; Gaps 48;				
QY	275	CEPCPGTY-----ADKSSFCCLC-----PANS-----YSN-----KGETSC-----	308	
Db	374	CEQCKPFYFQHPERDIPRNPICCEPCTCPAGSENGIDGDTDFSVGLAQCCKLIVE	433	
QY	309	-HQC-----PDYSEKSSGCV-----PACTDKDYF-----THTACDA	344	
Db	434	GERCDVCKEFGYDLAEDPYGCK-SCACNPIGTIGPNPCDSEGTGYCKRLVTGQRCD	491	
QY	345	NGETOLMKKAPKICSEDLGAVKLPASGVKTHCPNCPGFFKTNSTCQ-----PCRYG	400	
Db	492	---QCLPQHWG-----LSNDLQ-----CRPCDCLGALNNSCSEDSQCCL	532	
QY	401	SYNSGSDCT-----RCPAGTEPAVGF-----	421	
Db	533	PHMIGRQCNVESEGYFTTLTHIYEAEFANIGPCVVYVEROYIDRIPLSWTGP--GVYR	590	
QY	422	-----EYKMWNLPTMETTVLSGINFYKGMTGWEVAGDH1YTAGASDNDEMIITL	474	
Db	591	VPEGAYLEF-FIDNPYSMEYEL--IYEPQLPDHWEKA-----VITV	631	
QY	475	VPGFRPQOS-----VMAETENKEV-----ARTTFYELCSNCLYMGVY--SRINT	523	
Db	632	QRPGRIPASSCGMTVPDDNQVYSLSPGSRVYLPREVC-----FEKGNYTVREL	684	
QY	524	PVEFMKSGSKOSYTYLE-----ENTTSFTW-AFORTFHEASR	563	
Db	685	PQYASGSDVSPYTFIDSLVLMPCSKSLDFTYVSGSDGEVTNAMESMTPFQRIENSR	744	
QY	564	KYT-----NDVAK--IYSINVTVMNGVASYCRP-----CALEASDVSSCT	603	
Db	745	SVVTPMTDVCNRNIIFFSALIHQGLACEDDPQSLSVCDPNCGQOCQCRPNVYGRICN	804	
QY	604	SCPAG-----YIIDRSGTCH-----SCPNTTLKHAQ	631	
Db	805	RCAGTGFEGFGNGCKPCDCHLQGSASAFCDATGQCHCFQGIYARQCDRCILPGY-----	858	
QY	632	PYGVQACVPCGPRGTNNKIHSLCYNDCTPFSNRTPTRTNYNFSALANTVTLAAGSPFSK	691	
Db	859	-WGPSCQPC-----QCGNHL--DC-----DYVT--GCCLSCQ	887	
QY	692	GLKFFHNTLLSLCNGGRKMSVCTDNYDLRIPEG-ESG--FSKSI-----TAYVCO	740	
Db	888	DYTGHNENKRLAGYGRPIIGSGDHCRPCRPDRSDRGQFARBCYDDPYTLQACVCD	947	
QY	741	AVIIIPREVTGKAGVSSQVSLADRLIGVTTDMTIDGITSFAELFNLISLIPDY-----	795	
Db	948	-----PGY-----IGSRCDCCASGFFGNPSD--FGSGCQPCQCHNHNIDTTPREACDRDT	994	
QY	796	-----IIFYRRNDYQSGSSG-----RSTIIRVNCSPQKTVR-----GSL--LPQTC	836	
Db	995	GRCLKLYHTGDHCQLCYGYGDALQDCKCKVCNLTGVKEKNGSDCHCDKATQCD	1054	
QY	837	S-----DGTCDGCFHFLME--SAAACPLCSVADYHAIIVSSC	871	
Db	1055	SLCNLVGNQNDRCAPN-TWQLASGTGGPCNCAHAFSGESC	1096	
RESULT 15				
FUR2_DROME				
ID_FUR2_DROME				
STANDARD; PRT; 1680 AA.				

AC P30432;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Furin-like protease 2 precursor (EC 3.4.21.75) (Furin 2).
 GN FUR2.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92381036; PubMed=1512259;
 RA Roebroek A.J.M., Creemers J.W.M., Pauli I.G.L., Kurzik-Dumke U.,
 RA Reutrop M., Gattef E.A.F., Leunissen J.A.M., van de Ven W.J.;
 RA "Cloning and functional expression of Drosophila melanogaster with multiple
 RT proteolytic processing enzyme of Drosophila melanogaster with multiple
 RT repeats of a cysteine motif.";
 RL J. Biol. Chem. 267:17208-17215 (1992).
 CC -1- FUNCTION: FURIN IS LIKELY TO REPRESENT THE UBQUITOUS ENDOPEPTIDASE
 CC ACTIVITY WITHIN CONSTITUTIVE SECRETORY PATHWAYS AND CAPABLE OF
 CC CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Release of mature proteins from their
 CC propeptides by cleavage of Arg-Xaa-Yaa-Arg-|-Zaa bonds, where Xaa
 CC can be any amino acid and Yaa is Arg or Lys. Releases albumin,
 CC complement component C3 and von Willebrand factor from their
 CC respective precursors.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8. FURIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 DR EMBL: M94375; AAA28551.1; -;
 DR PIR: A43434; A43434.
 DR HSSP: O99405; 1MPT.
 DR MEROPS: S08.049; -;
 DR Flybase: FBgn004598; Fur2.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR002884; P-domain.
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR Pfam: PF01483; P; PARTIAL.
 DR PRINTS: PR00723; SUBTILISIN.
 DR ProDom: PD000717; P-domain; 1.
 DR SMART: SM00181; EGF. 1.
 DR SMART: SM00261; FU; 10.
 DR ProSITE: PS00136; SUBTILASE_ASP; 1.
 DR ProSITE: PS00137; SUBTILASE_HIS; 1.
 DR ProSITE: PS00138; SUBTILASE_SER; 1.
 KW Hydroxylase; Serine protease; Glycoprotein; Signal; Transmembrane;
 KW Multigene family; Zymogen; Repeat.
 FT SIGNAL 1
 FT PROPEP 1
 FT CHAIN 320 1680
 FT ACT_SITE 418 418
 FT ACT_SITE 457 457
 FT ACT_SITE 638 638
 FT DOMAIN 962 1444
 FT REPEAT 962 1007
 FT REPEAT 1008 1057
 FT REPEAT 1058 1104
 FT REPEAT 1105 1153
 FT REPEAT 1154 1205
 FT REPEAT 1206 1254
 FT REPEAT 1255 1299
 FT REPEAT 1300 1346

FT REPEAT 1347 1393
 FT REPEAT 1394 1444
 FT TRANSMEM 1508 1532
 FT DOMAIN 1533 1680
 FT CARBOHYD 3
 FT CARBOHYD 109 109
 FT CARBOHYD 130 130
 FT CARBOHYD 203 203
 FT CARBOHYD 443 443
 FT CARBOHYD 481 481
 FT CARBOHYD 928 928
 FT CARBOHYD 1061 1061
 FT CARBOHYD 1182 1182
 FT CARBOHYD 1275 1275
 FT CARBOHYD 1278 1278
 FT CARBOHYD 1440 1440
 FT SEQUENCE 1680 AA; 183599 MW; 0A99CE8770A8E293 CRC64;
 Query Match 3.1%; Score 169; DB 1; Length 1680;
 Best Local Similarity 17.2%; Pred. No. 0.00038;
 Matches 159; Conservative 85; Mismatches 291; Indels 388; Gaps 41;
 QY 46 LKAKSEHYHYETACDSG-----SRMY-----AVPHTPGCTSL 82
 DB 965 LHSCD-----AECDSGGYGRGPTQCVACSHYRLDNTCVSRCPSPRPMQVGIWCP 1016
 QY 83 PPVVKTECFSCNAGEFLMDKQSCPCAEGR---YSLGTGIRPEDEMDLPHGFASLSA 139
 DB 1017 HD-----TCETCAGAG-----PDSCLTCAPAHMLHYIDLAVCLQF-----CPDGYENSR 1060
 QY 140 NMELDSSAESTGNCSTSKWVRPGDYIAFNTDECATLTMAYANLKQSGVNFREYYPDS 199
 DB 1061 NR-----TCVPCBPNCASCODHP-----ECVTSK-----DNLVMEHKCYASC 1099
 QY 200 IIEFFPVNDQDQ-----PNAEDSRMKKTEKGMFHSVYELNCGNVLYVRTA 248
 DB 1100 PLDTYETENKCAFCHSTGATCNGPTDQCTICRSRRATWQ-----NKLISCPRG 1150
 QY 249 FSWTKVPKPVLVARNIAITGVAITYSECPCKPGYADKQSSFCPLKPNASYSNKGSTSC 308
 DB 1151 FYADKK-----RLECMPCQGS-----CKTCTSGNV-----C 1176
 QY 309 HOC-----DPDKYSEKSSSCSNVRPACTDQYFT-----HTACDANGETQLMY 352
 DB 1177 SECLQWTLNKRDKICVSGSEGS-----ESEFSQVEGOCRPCHASC----- 1219
 QY 353 KNAKPRICSEDEGANKPLASGVKTHCP-----CNPGEFTNNNSGOCPCFY- 399
 DB 1220 -----GSCNGPADTSCSPNRLLEDSRCVSGCREGFVEAGSLCSPCLHT 1266
 QY 400 -GSYNSDCTRCAPAGTEPAVGFEYKWMNTLPTNMETVLSGINFYKGMTGMEVAGDHI 458
 DB 1267 CSQCVSRKSNCSKGLQLONG-----ECITTADGY-YSDRIGACKCYLSCH- 1313
 QY 459 YTAAGASDNDEMLTLVDFGRPRPOSVMADTENKEVATITFVEPLCSVNCCLYFVAGVN 518
 DB 1314 -TCSGPRRQ-----ECITTADGY-YSDRIGACKCYLSCH- 1335
 QY 519 S-RTNPTVTKWGSKGKQSYTYIIEENTTSFTWAFQRTTPHASKRYNDVAKIYSIV 577
 DB 1336 ECHRECPGEGFYSDRCQCKH----- 1357
 QY 578 TWMNVGASVYCRPCALEADVGS-SCTSCTP-----AGYUIDRDSGTCNSCP 622
 DB 1358 -----YCKTC-----NDAGRIACLTSCPRSHMLDGLCMCLSSQYIDTTSATKTC- 1403
 QY 623 PWTILKAHPYGVQACVPCGPGCTKNNKHSCLYNDG---TFSRNPPTRTFFNFSALAN- 678
 DB 1404 HDSRSCFGP-GQFSGCKGCVPLHLDLQNSQVSCQONQTLAEKITS-----AACSC 1455
 QY 679 -----TYTLAGGPGFTGKGLKYLFIHFTLSLQ---NGGRKMSVCTDWTDLRIPEGS 728
 DB 1456 DGETGCKATSTGKRRITVGGSAVKSSESKHGSFENDGNAREV-----LRDLSPLT 1509

